Essays on the Economics of Education

by

Rigissa Megalokonomou

Thesis

Submitted to the University of Warwick

for the degree of

Doctor of Philosophy

Department of Economics

September 2015

THE UNIVERSITY OF
WARWICK

Contents

List of Tables

List of Figures

Acknowledgments

I wish to thank my supervisors Prof. Victor Lavy and Prof. Fabian Waldinger for their continuous support, their general guidance and for having their faith in me.

I am greatly indebted to Prof. Lavy and Prof. Waldinger who kept me motivated all the time and their meticulous comments were very crucial for completing this thesis.

Working closely to Prof. Lavy, for a project not included in this Thesis taught me how ideas grow up. His unwavering commitment to the highest standards of scientific work clearly permeates his entire lifestyle.

I am also thankful to Prof. Sascha Becker and Prof. Simon Burgess for helpful discussion and comments at my thesis defence.

Chapter 2 is based on my Job Market paper. I am particularly grateful to Prof. Waldinger for encouraging me to implement a large scale primary data collection for the accomplishment of this paper. The idea for this project was born during the PhD meetings that took place during my third and forth years in Warwick and were organised by Prof. Waldinger and Prof. Becker. I am particularly grateful to them for our discussions, their suggestions and most importantly for making me a better thinker.

I have also benefited a lot from discussions with Dan Bernhard, David Card, Clement Chaisemartin, Micro Draca, Caroline Hoxby, Kala Krishna, Robin Naylor and Roland Rathelot. Chapter 1 has benefited from various presentations at the University of Warwick in 2013, 2014 and 2015 (CWIP and CAGE AMES) but also from presentations at: 30th Annual Congress of the European Economic Association at the University of Manheim, XXIV Meeting of the Economics of Education Association at the Autonoma University in Madrid, 2nd Applied Econometrics IAAE at the University of Macedonia in Thessaloniki, AUEB Conferences in Economics and Econometrics in Crete and Milos, 18th IZA European Summer School in Labor Economics in Amersee, LEER Workshop on Education Economics at the University of Leuven, 2015 Royal Economic Society Conference at the University of Manchester and RSS seminar talk in Essex.

I am also grateful to Prof. Yanis Varoufakis for having inspired me back in my bachelor years to pursue a PhD.

I am always grateful to my family for keeping me grounded with their unconditional love and unfaltering support. They walked the way with me sharing the peaks as well as the valleys.

Funding from Varvaresou Legacy, University of Athens is gratefully acknowledged.

Statement regarding the Dataset

I am the only holder of the administrative data kindly provided by the Hellenic Ministry of Education. I am thankful to the Hellenic Independent Authority of Personal Data for giving a positive recommendation regarding my request. I would also like to thank Prof. Heracles Polemarchakis for helping me obtain the administrative data.

The school level dataset was collected together with my co-author Sofoklis Goulas who is a PhD student at the University of North Carolina at Chapel Hill. It took us two years to collect this dataset. I have visited more than 200 senior High schools in Athens, North areas in Greece, islands and rural areas. That was the most adventurous activity I undertook during my PhD.

Declarations

I hereby declare that the work in this thesis was carried out in accordance with the Regulations of the University of Warwick. I have not used before or published any material contained in this thesis, and it has not been submitted for a degree at another university

The papers related to the Chapters are joint work with Sofoklis Goulas- a PhD student at the University of North Carolina at Chapel Hill. I contributed more than 50% to the analyses and the content of the second chapter. I contributed equal to 50% to the analyses and the content of the third chapter. I contributed less than 50% to the analyses and the content of the fourth chapter.

Chapter 1

Introduction

Nowadays and more generally, discrepancies in economic growth between otherwise similar countries are vast and in a large extent unexplained by economic theory. Economists in their endeavour of disentangling this puzzle bring education in the frontline as empirical evidence indicates that in some cases reforms in education are significant in explaining shifts in economic performance. This thesis consists of three papers which address different questions in related fields regarding the economics of education.

The second chapter of this thesis concerns the effect of releasing information to students about their relative performance within their school and nationwide. Knowing how one's characteristics compare to those of other individuals is important in every setting of economic decision making. This chapter examines the effects of providing relative performance information on students' short and long term outcomes. I exploit a large scale natural experiment that took place in Greece. Using unique primary data on students' performance throughout senior high school, we find an asymmetric response to feedback: high-achieving students improve their final-year performance by 0.15 of a standard deviation, whereas the final-year performance of low-achieving students drops by 0.3 of a standard deviation. The results are consistently more pronounced for females indicating greater sensitivity to feedback. I also document the long-term effects of feedback: high-achieving students reduce their repetition rate for the national exams; they enrol into university departments that are more selective by 0.15 of a standard deviation and their expected annual earnings increase by 0.17 of a standard deviation. By contrast, the results for low-achieving students are negative. I provide suggestive evidence that feedback encourages students from low-income neighbourhoods to enrol in university and to study in higher-quality programs, which may, in the long run, reduce income inequality.

The third chapter of this thesis examines the extent to which college decisions among adolescents depend on the decisions of their peers. In the recent years, the importance of one's group of peers-be that friends, colleagues, neighbors- has been widely emphasized in the literature. In this paper, I ask whether individuals derive utility from conformity in college enrolment. I propose a new methodology in mitigating reflection and endogeneity issues in identifying social interactions. The instrument that I propose is the percentage of females in one's school, neighbourhood and preferacture the year before. Evidence from the psychology literature support our assumption that the prevalence of females creates a less violent and disruptive environment. I exploit a special institutional setting, in which schools are very close to each other, allowing for students from different schools to interact. I investigate utility spillovers from the educational choices of students in consecutive cohorts. Spatial variation allows us to identify social interactions in groups of various sizes, using a new dataset that spans the universe of high school graduates. I find positive and significant externalities in the decision to enrol in college among peers who belong to the same social group. Results indicate that students who attend high school with 10% more classmates who enrol in college are 4.5 % percentage points more likely to themselves attend college.

In the forth chapter, I investigate the causal effect of school attendance on students' performance. I exploit a natural experiment that changed the school absences allowance for the high achieving students in order to identify the effect of school attendance on educational outcomes. I use a novel dataset that contains class attendance information about students in eleventh and twelfth grade. The natural experiment took place in Greece in 2007 and provided higher performing students with 50 more hours of excused absences from school. I start off by using a Regression Discontinuity approach in order to measure the change in total absences and exam score due to the reform around the cut-off. The regression discontinuity cannot find an effect around the cut-off. The reason behind that is that the effect might not be caused by students around the threshold but by students in the right tail of the performance distribution. Next, I employ a combination of differences-in-differences and instrumental variables techniques in order to identify returns to absences. Our estimates show significant negative returns to absences.

Chapter 2

Knowing who you are: The Effect of Feedback Information on Short and Long Term Outcomes

2.1 Introduction

Improving pupils' attainments has been an important issue for policy makers and academics alike. In an effort to improve students' grades, education policies have focused on improving school inputs such as reducing class size (Angrist and Lavy 1999, Krueger 1999), improving the quality of teachers (Chetty et al. 2014, Rothstein 2010, Aaronson et al. 2007), extending the term length (Card and Krueger 1992) and improving the quality of the peer group a student is exposed to (Lavy et al. 2012, Zimmerman 2003a, Hoxby 2000b). All these interventions are significantly more costly than manipulating the availability of social comparison information. However, little is known about whether providing social comparison information enhances students' performance.

This paper presents a theoretical motivation and empirical analysis of whether providing high school students with social comparison information regarding their performance in externally marked high stake exams affects future performance in similar exams. Our analysis relies on the fact that different cohorts have different policies regarding the provision of feedback. The feedback policies we observe differ based on whether students receive information about their ordinal rank position at the end of the eleventh grade.

We exploit a large scale natural experiment that took place in Greece in 2005. Until 2005, all students were provided with relative performance information in a series of externally graded national exams prior to University admission high stake exams. In this regime, all students had to take national exams in two adjacent grades; one year before graduation from high-school and the year they graduated from high school (feedback regime). This system allowed students to receive information about their relative performance in the penultimate year. Knowing their performance in the penultimate year exams, students could translate their hours of effort into exam result. In the feedback regime, each student's performance in the eleventh grade exams was publicly announced, giving students the opportunity to calculate their national and school rank. So students could compare themselves to others allowing for social comparison. Knowing their relative performance in the eleventh grade could affect the amount of effort students decide to exert towards their twelfth grade performance (Ertac 2005). Students' performance in the twelve grade national exams is the most important determinant for University admission in this setting.

After 2005, the penultimate year national exams are abolished and replaced by school exams. This means that after 2005, penultimate year students sit exams on the same subjects as before but they now receive report cards with their own grades only. As a consequence, they no longer receive information about their penultimate year relative performance. These cohorts -as the previous ones- sit national exams in the twelfth grade that will determine their post-secondary placement. However, they have been imposed a loss of feedback information regarding previous performance in similar exam (non-feedback regime).

Using new data on school performance, school quality and national exams for university admission, we test the hypothesis that students' final year exam performance is independent of the feedback regime. Conditional on their tenth grade performance, we compare the final year performance of students across feedback regimes. After controlling for students' characteristics, we identify the effect of feedback provision on their short term (academic performance in the University entrance exams) and their long term outcomes (repetition of national exams one year after graduation, popularity of University Department admitted to and expected annual earnings).

Our first finding is that high achieving students perform better in externally graded exams when they are aware of their relative performance in the school and nationwide. Feedback information on past performance improves the next period's exam performance of the better students by 0.2 standard deviations and their relative national rank by 4-6 percentiles. This is of comparable magnitude to being taught by a teacher 1.5-2 standard deviations above the average (Chetty et al. 2014, Hanushek et al. 2005) or to reducing the class size by 15 percent. (Angrist and Lavy 1999, Krueger 1999). Additionally, we find evidence that the performance of students in the lower percentiles deteriorates when feedback is provided. In particular, their consecutive year performance declines by 0.3 standard deviations and their national rank decreases by 6-8 percentiles. To build intuition here, we consider that knowing how someone performs relatively to others in the same task affects his motivation to exert less or further effort in the next time period. (Ertac 2005)

Our second finding suggests differential response to feedback at different parts of the ability distribution by gender. Females seem to be considerably more sensitive to feedback at all parts of the ability distribution than males. In particular, high achieving males and females respond positively to positive feedback whereas low achieving male and female students respond negatively to negative feedback. Our results are consistent with the existing literature findings regarding gender differential response to feedback information due to initial different levels of self-confidence (McCarty 1986).

Our third finding is that the provision of feedback changes the matching of students to University Departments. After ranking all University Departments based on their popularity and cut-offs, we find that feedback provision makes high (low) achieving students move up (down) the University Departments popularity ladder by 30 (35) programs which is 0.15 (0.18) of a standard deviation. We then map each program with the annual earnings of older graduates. We find that when feedback is provided, high (low) achieving students experience an increase (decrease) in the expected earning by 0.13 (0.23) standard deviations. In absolute terms, we find evidence that feedback alters the socio-economic background composition of students who manage to get admitted to the top programs. More students from low income neighbourhoods get admitted to the most selective programs with the highest expected earnings after graduation (like engineering and law), when feedback information is provided. This means that feedback information encourages social elevation motivated by students from low income families.

To the best of our knowledge this is the first large scale study that documents the short and long term effects of feedback provision using a natural experiment. In particular, we document the direction and size of the effect of feedback information on students' attainments, post-secondary placement and expected earnings. We exploit a special setting where high school students receive information about their relative position in two reference groups (school and country). Thus, we define feedback as the information of one's performance in comparison to their peers in school and nationwide.

We also discuss the two most prevailing mechanisms that could explain how students react to the social comparison information and are related to students: 1) learning about own relative ability and/or 2) learning about the quality of the school. The mechanism that best accommodates all our findings is the first one with the second helping us rule out alternative interpretations of the results. This approach may seem as a departure from the usual behavioural formula, where individuals are only uncertain about other agents' type. Although someone may observe his own perspicacity, they do not generally observe everyone's performance so as to deduce a useful assessment of their own relative performance. We provide evidence that feedback provision can change good allocation in such environments.

In the recent years there has been an increasing interest in the economic literature of feedback information provision on exam performance.[1] Bandiera et al.

[1] The relative feedback information has been studied in the tournament literature. Some studies find that relative performance information has a positive effect for all participants in tournaments and piece rate payment schemes (Hannan et al. 2008). On the other hand, some other studies find mixing results. Barankay 2012 uses data on furniture salesmen's effort and finds that feedback has negative effects on the low performing employees.)

2008 examine the effect of feedback information on students' future absolute performance using data for University students registered to Departments with different feedback policies. In that study, feedback is defined as the knowledge of someone's absolute performance in the midterm exam in period one and before students exert effort on their essay in period two. The authors find that the effect of feedback is positive for all students and more pronounced for more able students. However, their study refers to feedback involving own performance. The provision of feedback regarding relative performance has not received much attention.

The paper which is most closely related to ours is a study by Azmat and Iriberri 2010. The authors examine the effect of relative performance feedback on students future absolute performance. They exploit a natural experiment that took place in a high school, where for one year only students received information about the average class score in addition to their own performance. Their findings suggest that feedback improves the performance of all students in the subsequent test. hey also find no differential effect by gender along the ability distribution. A key difference to our work is that they use a small sample of one high school while we use a sample of 134 senior high schools nationally representative in many dimensions. Another important difference is that Azmat and Iriberri 2010 investigate the effect of providing information about someone's class rank. We contribute to the literature by examining the effects of providing broader social comparison information about someone's school and national rank. More recently and in a related literature, Murphy and Weinhardt 2014 examine the effect of knowing one's ordinal rank position in exam results on future exam performance. They find large and robust effects of being highly ranked in primary school on secondary school achievement. Their study also reports that boys are more affected by knowing their ordinal rank than girls. Students figure out their rank within their class from social interaction with their classmates. In our setting, the information set is greater and is provided by the principal. Students receive explicit information regarding their rank position within the school and nationwide which facilitates the policy recommendations.

The chapter is organized as follows. Section 2 presents a theoretical model for the individual's behaviour and motivates the empirical investigation. Section 3 provides a brief description of the institutional setting and the data. Section 4 sets out our empirical strategy. Section 5 presents the main results on short and long term outcomes and discusses heterogeneous feedback effects by ability, gender, track and neighbourhood income. Section 6 discusses the threats to identification and reports further robustness checks. Finally, in Section 7 we conclude and discuss possible policy implications.

2.2 Theoretical Framework

We adapt a theoretical model proposed by Ertac 2005 [2] where students have imperfect information about their own ability.

In the non-feedback regime eleventh graders take school exams and they receive information about their own performance only. In the feedback regime, they receive information about their own performance and about the school and cohort average performance. Students engage in a task in two time periods; the eleventh and the twelfth grade. Students' performance in the eleventh grade depends on their ability and the easiness of the task. This performance provides them with some information about ability and easiness of the exam [3]; we will refer to that as the private signal s_i. The ability of a student is denoted by $\alpha_i > 0$ and α_i's are independent draws from the same distributions and independent of task difficulty. All distributions are common knowledge.

When signals coming from the eleventh grade are realized, students update their beliefs about the ability and decide the subsequent effort. The amount of effort students decide to exert in the twelfth grade determines their final year's scores. Period 1: This is the learning stage. Students receive a noisy signal about their ability:

$$s_i = \alpha_i + n \quad , i = 1, 2, ...$$

$$\alpha_i \sim N(\bar{\alpha}, \sigma^2), \bar{\alpha} > 0$$

$$n \sim N(0, \psi^2)$$

$$cov(\alpha_i, n) = 0, \quad cov(\alpha_1, \alpha_2 = 0)$$

This signal (s_i) depends on student's i ability level (α_i) and a shock (n) that is common to all students [4]ie. the easiness of the exam. We also assume that α_i and n are normally distributed and α_i and s_i are jointly normally distributed. In

[2] Ertac 2005 presents a principal-multiple agents model where agents have imperfect information about their abilities under multiple types of contracts. The model is also used by Azmat and Iriberri 2010. The natural experiment they study gives students information about the average grade of the class, while here the social comparison information refers to the average school and cohort grade.

[3] In the feedback regime n refers to the easiness of the national exams. In the non-feedback regime n refers to the easiness of the school exam.

[4] In the school or the cohort depending on the feedback or non-feedback regime.

the feedback regime, students also observe the average signal:

$$\bar{s} = \frac{\sum_{i=1}^{N} s_i}{N} = \frac{\sum_{i=1}^{N}(\alpha_i + n)}{N} = \frac{\sum_{i=1}^{N}(\alpha_i)}{N} + n$$

The type of the signal each student receives, affects student's perceived belief about his own ability in the first period. A student's belief about his own ability determines the amount of effort he chooses to exert in the second period. Then we find student's i expectation of his own ability conditional on the signal he observes in each case [5]. In the non-feedback regime the student observes his own performance in the school exams. His expected ability given the observed signal is:

$$E(\alpha_i|s_i) = \bar{\alpha} + \frac{\sigma^2(s_i - \bar{\alpha})}{\sigma^2 + \psi^2}$$

In the feedback regime, the student receives information about the average performance in his school but also nationally:

$$E(\alpha_i|s_i, \bar{s}) = \bar{\alpha} + \frac{(\sigma^2 + \psi^2)(s_i - \bar{\alpha}) - \psi^2(\bar{s} - \bar{\alpha})}{\sigma^2 + 2\psi^2}$$

The higher the private signal a student receives, the higher is his belief about perceived ability. If the average signal is observed, then the belief about ability decreases with it.

Period 2: Following the realisation of the signals, students choose in the second period the amount of effort to exert. Notice here that there is no pass-fail scheme and students do not try to achieve a specific performance threshold. University cut-offs are determined endogenously based on demand and pre-specified supply of seats. In other words, we assume that ability and effort are complements in the production function. [6] Assuming that the performance production is a linear function in effort [7] and that effort and ability are complements in performance[8] it follows that: $q_i = \ e_i\alpha_i$. There is also a cost associated with the effort exerted that is

[5]Using the properties of the multivariate normal distribution, we find that
$$\begin{pmatrix}\alpha_i\\ s_i\\ \bar{s}\end{pmatrix} \sim \begin{pmatrix}\bar{\alpha_i}\\ \bar{\alpha_i}\\ \bar{\alpha_i}\end{pmatrix}\begin{pmatrix}\sigma^2 & \sigma^2 & \sigma^2/N\\ \sigma^2 & \sigma^2+\psi^2 & (\sigma^2+N\psi^2)/N\\ \sigma^2/N & (\sigma^2+N\psi^2)/N & (\sigma^2+N\psi^2)/N\end{pmatrix}$$

[6]In a different setting where university cut-offs are pre-determined, effort and ability could be substitutes in the production function. In that case, a student who is above average in the eleventh grade may choose to exert less effort in the twelfth grade in order to achieve a specific performance threshold.

[7] The predictions of the model do not change if the performance function is not linear in effort. But we make this assumption here for simplicity.

[8] $\frac{dqi}{d\alpha_i de_i} > 0$

$c(e_i)$ and is increasing in effort and convex. [9] In the second period, students choose the effort level $e_i > 0$ in order to maximise their last year's utility function. In the absence of feedback students receive only the private signal and they maximise:

$$u^{NF} = E[q_i - c(e_i)|s_i] = E[\alpha_i|s_i]e_i - c(e_i)$$

and the F.O.C simplifies to $E[\alpha_i|s_i] - c'(e_i^{NF*}) = 0$ (1)

In the feedback regime each student observes the average signal and maximises:

$$u^F = E[q_i - c(e_i)|s_i, \bar{s}] = E[\alpha_i|s_i, \bar{s}]e_i - c(e_i)$$

and the F.O.C simplifies to $E[\alpha_i|s_i, \bar{s}] - c'(e_i^{F*}) = 0$ (2)

Given that the F.O.Cs are sufficient, we will compare the optimal effort levels in the two regimes. The conditional expectation of ability is independent of effort while the second term in (1) and (2) is an increasing function of effort. That means that an increase in the beliefs about ability -a higher self confidence level- leads to an increase in the optimal effort level. The comparison of the F.O.Cs for the two regimes simplifies to the comparison of the conditional expected abilities.

$$E[\alpha_i|s_i, \bar{s}] = E[\alpha_i|s_i] \quad if \quad s^* = (\bar{s} - \bar{\alpha})\frac{N(\sigma^2 + \psi^2)}{\sigma^2 + N\psi^2} + \bar{\alpha}$$

Thus, if $s_i > s^*$ then $e^{F*} > e^{NF*}$ and if $s_i < s^*$ then $e^{F*} < e^{NF*}$. [10]. Students with signal above (below) s^* will put in more (less) effort, when feedback is provided. If $\bar{s} = \bar{\alpha}$ then the exam is neither hard nor easy. If $s^* = \bar{\alpha}$ which means that $s^* = \bar{s}$ and the average signal equals the average ability level and $e^{F*} = e^{NF*}$. However, if $s^* > \bar{\alpha}$ then $s^* > \bar{s}$ and if $s^* < \bar{\alpha}$ then $s^* < \bar{s}$. That means that if the signal is above the average signal then students will exert more effort when feedback is provided. Similarly, if the signal is below the average signal then students will exert less effort when feedback is provided.

If $\bar{s} > \bar{\alpha}$ then the exam was hard and the signal needed in order for students to exert more effort is higher than the average signal($s^* > \bar{s}$). If $\bar{s} < \bar{\alpha}$ then the exam was easy and the signal needed in order for students to exert more effort is lower than the average signal($s^* < \bar{s}$)

Let us summarize now the main hypothesis about the effect of the eleventh grade social comparison information on the twelfth grade performance.

[9] $(c'(e_i) > 0, c''(e_i) > 0, c'(0) = c''(0) = 0)$

[10] $\frac{N(\sigma^2+\psi^2)}{\sigma^2+N\psi^2} > 1$ provided than $N >= 2$

Null Hypothesis: Students do not react to the social comparison information

That would suggest that students are not uncertain about their ability or that students have already figured out their relative performance information and the explicit addition of it is redundant or that the private signal that students get in the feedback regime equals the average signal.

Alternative Hypothesis: Positive effect on performance for high ability students and negative effect on performance for low ability students

That would suggest that students will react differently to feedback. Based on the model, high ability students will perform better when the social comparison information is provided because they have been encouraged by their period one performance. On the other hand, low ability students will performance worse when the social comparison information is provided because they have been discouraged by their eleventh grade performance.

2.3 Institutional Setting and Data

2.3.1 Institutional Setting

In Greece, all students in secondary education are obliged to take the national exams to have access to tertiary education. Students sit these national exams in specific subjects on specific dates every year and the questions asked are the same for all students across the country. All universities are public and the admission procedure is run exclusively by the Ministry of Education. The University admission in Greece is based on the "admission grade". The admission grade is a weighted average of the grades a student gets in the national exams (70% weight) and the school grades (30% weight). The school grade of every subject is the average of the term grades. Only final year students can participate in the university admission procedure. Admission is made in a specific university department. All students are examined on the General Education core modules. On the top of that, students are examined on Elective subjects that are determined by the "speciality" or the "track" they choose at the beginning of the twelfth year.

After the admission grades are announced, every student makes and submits to the Ministry of Education a preference list of university departments he would like to be admitted to in that year. If a student is admitted to a school in a higher place in his preference list he cannot be admitted to those below that. That makes students to be very careful in constructing their preference list. The only way a student can flee from the university admission procedure is to deny submitting a list of preferences. Every university department admits a pre-specified number of students. Then, each department admits the best students that have included this department in their preference list. All students are compared to each other according to their admission grades and every successful candidate is admitted to the first department in his list where there is an available place and every student with higher admission grade has already been allocated. The rest of the students are denied admission at that year.

At the end, every department announces the grade of the last student it admitted in that year. This grade is considered to be the "bottom grade" or the "cut-off grade" in that year of a university department. More popular departments exhibit higher bottom grades. Students are aware of the " cut-off grades" of the previous years. The ranking of university departments according to their cut-off grades appears to stay largely unchanged, year after year, and this represents the society's valuation for these departments. It's not possible to defer someone's admission. Some students that have not been admitted to the university department they wanted to may decide to retry admission a year (or more) after graduation using their school grades in the admission procedure and retaking national exams in all subjects. Those students usually do not attend any school/college or pursue any job or do military service after graduation and before the next admission period.

2.3.2 Data Collection

The transition from high school to higher education is based on a centralised allocation of students to University Departments. The Hellenic Ministry of Education collects data on all students on the years that they sit national exams. The performance of students in previous grades can only be found in the archives of the school they attended. Thus, we visited senior High schools across the country and we have constructed a database of detailed student performance in every subject throughout senior high school. In particular, we have collected data from a large randomised sample of 134 schools across the country. Our novel dataset combines information from various sources:

1. Administrative data provided by the Ministry of Education regarding the twelfth grade performance of all students who sat the twelfth grade national exams from 2003 to 2009. This dataset contains student level information about gender, national and school exam results in each subject nationally examined in twelfth grade, name of senior High school attended, year of birth and

graduation year from senior High school, speciality chosen at the beginning of twelfth grade. It also contains University admission related information such as the University Department each student got admitted to, number of applications made to University Departments and the reported ordinal preference position of the University Department admitted in someone's preference list. The dataset refers to the period 2003-2009 and gives us information about 435.589 students.

2. As the Ministry does not collect information on students' tenth grade performance, we collected this information directly from the schools.[11] More specifically, we have physically collected data from 134[12] public, experimental [13] and private schools both near big cities and in the countryside (this number corresponds to around 10 % of the school population). We exclude the evening schools[14] from our analysis because they differ in many aspects from the other types of schools. [15] This dataset includes information about school and/or national exam results in tenth, eleventh and twelfth grade in all subjects, indicators for gender, a class indicator, graduation year, year of birth, speciality chosen at the beginning of the eleventh and twelfth grade and a unique identification code for each student that stays the same throughout senior high school. We have had short interviews with the principal of every school in our sample to find out about any effects potentially affecting our outcomes of interest. Inter alia, principals were asked about the size and history of the school, facilities, attrition and teacher quality. The matching between the dataset provided by the Ministry of Education and the school datasets was almost perfect [16] providing us with a complete senior high school performance history for 45.746 students which is our sample size.

3. The Ministry of Economy and Finance provided us with average household income information for 2009 for every postcode in the country. We employ this as a proxy for neighbourhood income.

[11]The tenth grade performance data are recorded in each school's archives either in their computers or in their history books.

[12] We exclude from the analysis schools that had at least one year school cohort size smaller than ten students because these small schools may be atypical in some dimensions.

[13]Experimental schools are public schools where admission in these schools is based on a randomised lottery.

[14]Which are public schools but lessons take place in the evening targeting employed students.

[15]University cut-offs differ for students graduating from evening schools compared to any other type of school.

[16] 92 % of students matched because of missing values regarding the year of birth or the gender of the student in the school data.

4. The Ministry of Internal Affairs provided us with urban density information. Urban areas are those with more that 20,000 inhabitants.

5. The National Statistical Authority provided us with the Labour Force Survey data for the year 2003. We use quarterly data to create a variable that maps college occupations into annual earnings [17]. Respondents report their occupation with high precision.[18] The earnings data are grouped into ten bins indicating the ten national deciles with the highest frequency. We use the lowest bound of each bin [19] to construct a variable that measures minimum expected annual earnings from each occupation.

Every school follows the same curriculum and students are assigned to public schools based on a school district system. This school district system assigns students to schools based on geographical distance. Students are alphabetically assigned to classes in tenth grade and then they do not change class throughout senior high school. Moreover, teachers are allocated to public schools based on geographical criteria and no quality criteria are taken into consideration in the process. Figure 2.1 presents the geographic position of each school included in the sample. The density of the school population in Athens is 32 % thus many of the schools in our sample are located in Athens.

Table 2.1 presents descriptive statistics about the available variables in the sample in the twelfth grade. The variable "internal migration" takes the value one if the district of University Department the student is admitted to is different from the district of residence; the latter being proxied by the school district. Moreover, the variable "early enrolment" takes the value of one if the student enrols in the first grade before the age of six [20]. Interestingly, 82 % of the students on average get admitted to at least one University Department. Given that there are no fixed cut-offs, if there is not much demand for a particular University Department the cut-off grade in that year is very low.

Table 2.2 reports the mean characteristics of the schools in our sample and the whole school population to investigate if our sample is a representative one. There are some variables for which there is a statistically significant difference between the 134 sample schools and the population of schools and these differences are mainly

[17] We also map college fields to occupations.

[18] 209 classified occupations are reported and respondent have to indicate which one is closest to their actual one.

[19] Multiplied by 12 months.

[20] According to the law, this happens if the student is born in the first quarter of the calendar year.

related to the sampling methods that we use [21]. So the sample may not be fully representative of national responses, but it looks pretty similar nonetheless.

2.3.3 How does feedback work?

Knowing one's own relative performance might affect the amount of effort a student exerts with regard to a certain objective. In the context of our study, the student's objective is to maximize his or her score and/or rank at the end of high school.

Consider a student in the treated group. In the world of this experiment, students compete with each other over access to a limited number of university places. At the end of the penultimate year, students take standardized exams in some subjects with external examiners and at least two anonymous external graders per subject.

Then two mechanisms are in action. First, everyone's results within the school become public knowledge: the names and detailed grades are displayed at the entrance of every school. This provides students with information about how well they can do given a specific level of effort, when national exams come around again. This means that students could calculate their distance from the school's average score, and their relative position within their school. Second, the names, details about national exam scores, and the cohort's average national exam score are published in the newspaper. This means that each student could calculate her distance from the national cohort's average score and derive her relative national rank. The students' names and their grades are sorted based on an alphabetical order.

We believe that students in the feedback regime calculate their eleventh grade rank within the school and nationally given the importance of their performance in the senior year exams. Knowing a student's national rank provides them with information about the competition in that year. Each year the newspaper reports the following: cohort's average national exam score, the cohort's minimum and maximum score, the score that corresponds to each decile and comparisons with last year's statistics. For each student the following is reported: student's first name, surname and father's name, score given by the first and the second examiners (Figure 2.3 and 2.4). This is published separately for each subject. The score given by each examiner ranges from 0 to 100. If the difference between the score given by the first and the second examiners is not greater than 13/100, then the final score

[21] i.e. the relative percentage of schools in Athens for which we collected data is higher than the relative percentage of schools in Athens. Furthermore, our sample contains 5 % fewer private schools than the population.

is the average score between the scores given by the first and the second examiners. Otherwise, the final score is the average between the highest two scores given by any examiners. The raw final score used a 1-to-20 scale that we transformed into z-scores to facilitate the interpretation of the results.

Students use this information to calculate their national rank. Given that the names are alphabetically sorted, calculating a student's rank using even a single newspaper's scores is already a good indicator of a student's national rank. Calculating the school rank is much easier given that the average school cohort consists of 79 students.

Consider a student in the control group. During the penultimate year of senior high school, he chooses an effort level to prepare for his exams, which are now given only at the school level. Within the school, teachers coordinate to cover the same material, and usually give the same exam questions. Before the summer break in the penultimate year, our student takes exams on the same five subjects, and receives a written report from school with his own grades. When he reaches twelfth grade, he has access to the same material, study guides and past exam papers as any student in the treated group. However, he is unaware of how his schoolmates and his cohort did relative to him in the penultimate year final exams. Table 2.3 reports the summary statistics of the variables of interest across the two regimes. Some of the differences seem to be significantly different from zero but they are either very small or economically non-meaningful. The exact timing is presented in Figure 2.2.

2.3.4 Test Scores

Our prior measure of performance is based on the overall students' performance in the tenth grade (GPA). The tenth grade GPA takes into account a student's tenth grade performance in thirteen subjects. All tenth grade exams are school exams. The performance in each subject is a weighted average of the final exam result (50% weight) and the performance of a student during the school year (50% weight). Teachers receive guidance on how to mark the final tenth-grade school exam. We use the within-school rank of each student based on the tenth grade GPA as a prior measure of performance. It is compulsory for all students to take exams in all thirteen subjects in tenth grade.

Our main outcome variable is a student's twelfth grade rank in two reference groups (the school and nationally). These outcome variables -the within school and the national rank- take into account students' twelfth grade performance in five core-education subjects. Students take exams at the end of the twelfth grade in these core-education subjects. A student's performance in these five subjects is the most important determinant for the calculation of the high-school graduation grade under both regimes. Before 2005, these five subjects were all examined at a national level. From 2005 onwards, two subjects are examined at a national level whereas the other three subjects are examined at a school level. This change in the number of subjects examined at a national and school level happened in the same year as the abolition of feedback. We do various robustness checks later to examine if this change affects our results. In particular, we use various outcome variables (the rank in each subject separately; the average rank in those subjects examined at the national level; or the average rank in the five core-education subjects) and the estimated effects follow the same patterns. We call the core-education subjects "incentivized", because performance in these subjects is taken into account in the calculation of the admission grade.

All schools in the sample offer three academic tracks in the twelfth grade. Each student has to choose the academic track that is relevant to the post-secondary degree they desire to pursue. Each track offers different subjects. Depending on the track students choose, they take national exams in four track-specific subjects in both regimes [22] We do not include the test scores in these four subjects in the main analysis because the choice of track is based on endogenous criteria, i.e.their perceived differential ability or preferences for a particular degree after high school graduation. Robustness checks show that the results remain almost unchanged when the track-specific subjects are taken into account.

In addition to the core-education subjects and the track-specific subjects, students take compulsory within-school exams in three subjects (Sociology, Religion and Modern Greek Literature) in both grades; eleventh and twelfth. Students take school exams at the end of the eleventh grade and each student receives a report card. This report card shows each student's own performance in these exams without providing information about the class or school average score. In the twelfth grade, students are examined again on these subjects without having previously received any relative performance information in these three subjects. We call these subjects "non-incentivized" because students' performance in these subjects is not taken into account in the calculation of the university admission grade in any of the regimes. Students take these exams in both regimes. We use these subjects as the main counterfactual group.

In our analysis, we use the rankings instead of absolute scores for a couple of

[22] These four subjects differ from the one track to the other. The Tracks are: Classics, Exact Science and Information Technology.

reasons. First, using the tenth-grade ranking allows us to make comparisons across cohorts and across schools. Notice that we do not observe the different feedback policies in the same year. Thus, we use the within-school rank of a student to compare students who are exposed to different peer groups and teachers. Second, a given twelfth-grade national exam score does not represent the same ability level in different years. However, it is important to make sure that students of the same ability obtain the same rank in different years. The comparison of students' absolute scores across cohorts would be problematic, if the difficulty of the exam changes from one year to another. Additionally, calculating the ranking of a student in each subject takes into account the potential difference in the difficulty of the exam from one subject to another. Thus the ranking allows us to compare a students' performance across different subjects. Also note that any school grade inflation that might occur in the tenth grade does not affect our prior performance measure (tenth grade GPA). Grade inflation would make the teacher more lenient in the overall grading procedure, which implies that the ranking of the students remains unaffected. The national exams in twelfth grade are externally graded. As a result, the teacher in a student's school has no way to affect their national exam final scores. Furthermore, the national exam procedure does not receive any grade curving.

2.4 Empirical Strategy

This section identifies and magnifies the effect of relative performance information on students' exam performance. First , we define the rank measures that we use. Second, we identify if there is an effect. Since we use as an outcome variable the rank in the twelfth grade, the effect is -if anything- of a distributional nature. Then, we discuss the empirical method in order to identify the effect of feedback on students' last year relative performance.

2.4.1 Calculation of the rank

In order to calculate the relative rank of the student within his school in the tenth grade, we use the following normalization in order to allow comparisons across schools and cohorts:

$$Rank_{10isc} = \frac{n_{isc}-1}{N_{sc}-1}$$

where n_{isc} is the ordinal rank position of student i within school s in cohort c in tenth grade [23] and is increasing in GPA and N_{sc} is the school cohort size of school

[23] Based on the average of the thirteen subjects, ie.the tenth grade GPA.

s in cohort c. The higher the $Rank_{10isc}$, the higher the rank position of student i in tenth grade in his school s and cohort c. Moreover $Rank_{10isc}$ is bounded between between 0 and 1, with the lowest rank pupil in each school having $R_{10isc} = 0$.

The ranks of the student within his school in the twelfth grade and nationwide are calculated using the following normalisations:

$$Rank - school_{12isc} = \frac{k_{isc}-1}{K_{cs}-1}$$
$$Rank - nationwide_{12ic} = \frac{r_{ic}-1}{R_c-1}$$

Where k_{isc} is the ordinal rank position of student i in school s in cohort c in twelfth grade and is increasing in the national exam grade. K_{cs} is the cohort size c in school s. The $Rank - school_{12isc}$ is projected into the [0,1] interval and the lowest rank pupil in each school cohort has $Rank - school_{12isc}$=0. Notice that there are five exams/subjects, so we first find the ordinal rank of the student based on the average in the five scores, and then we normalise it using the above formula. $Rank - nationwide_{12ic}$ is calculated in a similar way but is irrespective of the school the student attends. So both $Rank - school_{12isc}$ and $Rank - nationwide_{12ic}$ are calculated based on the twelfth grade national exams in the incentivized subjects but they measure relative performance in the school and the country respectively.

2.4.2 Identifying the effect

Figure 2.5 shows the fitted values of the twelfth grade rank nationwide for each percentile of prior performance. We observe that the fitted regression line for the feedback period is steeper than the non-feedback one, implying that feedback has a positive effect on the better students and a negative effect on the students in the lower part of the ability distribution. Thus, the better students are more likely to end up higher in the twelfth grade rank distribution when feedback is provided. The opposite holds for the worse students who are more likely to end up lower in the twelfth grade rank distribution when they are aware of their previous relative performance.

In the same direction, Figure 2.6 shows the average rank nationwide of each performance group in the twelfth grade exams, conditional on students' prior performance. Cohorts up to 2005 have received feedback information and cohorts from 2006 onwards are used as a comparison group meaning that feedback is abolished. We observe that the lines are parallel in the treatment period (cohorts 2003, 2004 and 2005). This means that the time trends for each quintile of prior performance follow a similar pattern from year to year. Identification is achieved through a difference approach for each prior performance group. The 2006 cohort is the first cohort affected by the abolition of the relative performance information. We observe that from 2005 to 2006 the slopes of the time trends change, meaning that the treatment affected students in all performance groups considerably except the middle quintile, which remained unchanged. In particular, the top quintile achieved a higher average rank nationwide in the twelfth grade when feedback was provided compared to the period after 2006. The opposite applies to the bottom quintile where students end up lower in the distribution of twelfth grade rank when they are aware of their previous relative performance compared to the period after 2006.

Another important observation here, is that the slopes remain relatively stable after 2006, which is the first affected cohort. So, the change in the slope of the time trends between 2005 and 2006 can be attributed to the abolition of the relative performance information. We produce this figure using students' rank nationwide (Figure 2.6) and rank within the school (Figure 2.7).

2.4.3 Method

Here, we quantify the effect of feedback provision on future performance by adopting two complementary strategies.

First, we use the following specification to estimate the effect of feedback information on students' later rank, conditional on their prior performance.

$$Rank - nationwide_{12ic} = \alpha + \beta_{quintile} Feedback_c * Quintiles_{10isc}$$

$$+\lambda_{quintile} Quintiles_{10isc} + \psi Feedback_c + X'\gamma + \psi_c + \phi_s + \epsilon_{ic} \text{ (1a)}$$

$$Rank - school_{12isc} = \mu + \delta_{quintile} Feedback_c * Quintiles_{10isc} + \kappa_{quintile} Quintiles_{10isc}$$

$$+\xi Feedback_c + X'\zeta + \theta_c + \omega_{isc} \text{ (1b)}$$

where $Quintiles_{10isc}$ is a dummy variable that takes the value of one if the student is in the corresponding quintile based on his tenth grade performance in his school. Moreover, $Feedback_c$ is a dummy variable equal to one if the student takes the eleventh grade national exam ie. if the graduation year is smaller than 2006 (feedback regime). The parameter of interest β (δ) measures the effect of feedback on student's rank nationwide (within his school) in the subsequent year, conditional on tenth grade performance. In some specifications, we control for unobserved time and school invariant factors that may affect last year's rank using time and school fixed effects. Specification (1b) exploits within school variation, thus we use (1a) without the school fixed effects when we are interested in exploiting across schools time invariant variation.

In addition to the first strategy, we now use the following difference specification to find the effect of feedback on each decile of students' twelfth grade performance. We run the following specifications for each decile of tenth grade performance $\theta \in [0, 1]$:

$$Rank - nationwide_{12ic\theta} = \delta_\theta + \alpha_\theta X_{i\theta} + \beta_\theta D_{\mathrm{c}} + \psi_c + \epsilon_{ic\theta} \text{ (2a)}$$

$$Rank - school_{12isc\theta} = \omega_\theta + \alpha_\theta X_{ic\theta} + \gamma_\theta D_{\mathrm{c}} + \theta_c + u_{isc\theta} \text{ (2b)}$$

where δ_θ captures a performance group-specific fixed effect. The parameter of interest β is estimated separately for each one of the ten deciles, including clusters at the school level. A similar regression across all decile groups gives the pooled OLS estimator of β_θ which is exactly zero because as we explained before, the provision of feedback has a zero average effect. A negative coefficient of β_θ (γ_θ) implies that feedback induces a deterioration in the rank nationwide (within his school) for students at this decile.

2.5 Main Results

2.5.1 Effect on performance

Main OLS results are reported in Table 2.4. This table shows the effect of feedback on students' twelfth-grade national exam performance for each quintile of prior performance conditional on students' and schools' characteristics. The first column corresponds to the basic specification (1a) without school and year fixed effects. The dummy for the third tenth grade quintile is omitted as a point of comparison. This shows that when feedback is provided, a student in the top quintile in his school based on his tenth grade performance has a 0.042 percentile rank gain in his twelfth grade national exam performance compared to a student who is in the median quintile in his school, ceteris paribus. Similarly, a student who receives feedback and is in the bottom quintile in his school based on his tenth grade performance has a 0.088 percentile rank loss in his twelfth grade national performance compared to a student in the median quintile in his school. In columns 2 and 3, we see that the results in column 1 are robust when controlling for unobserved heterogeneity across schools and years respectively. Adding school and year fixed effects slightly change the coefficients estimates, which remain statistically significant at an 1 % significance level. In all specifications, we control for a set of pupil characteristics and we cluster the standard errors at the school level. These results support the alternative hypothesis of the model discussed earlier.

Female students seem to receive a lower national exam rank in the incentivized subjects. Additionally, students who are coming from high income neighbourhoods and urban schools perform better in the incentivized subjects. Public school is the omitted category regarding the school type. Students who attend experimental and private schools get a higher rank nationwide in the incentivized subjects compared to students in public schools. Furthermore, student who specialize in the science track experience a drop of 0.055 national percentile ranks in the incentivized subjects compared to students who specialise in computer science while students specialising in classics perform worse than students specialising in computer science.

We now turn to specification (1b) where we exploit the within school variation and results are presented in Table 2.5. The effect of feedback on students' within school performance in the incentivized subjects is reported in columns (1) and (2) and in the non-icentivized subjects in columns (3) and (4). In the first column, we show that students in the quintiles 5 and 4 (top ones) based on the tenth grade performance benefit from feedback. This gain is associated with 0.045 and 0.040 school percentile ranks respectively compared to the third quintile. Similarly, quintiles 2 and 1 (bottom ones) experience a loss of 0.038 and 0.079 school percentile ranks, when feedback is provided. In column 2 we control for unobserved heterogeneity across years and as we expect; results are similar to Table 2.4, column 3 when we conditioned out for unobserved heterogeneity across years and schools in the national analysis.

Then, we replicate the same analysis but we now use the school rank in the non-incentivized subjects as the outcome variable. As mentioned before, students take school exams in these subjects in both regimes and grades (eleventh and twelfth). This is a crucial placebo test because if students act as if they receive feedback in these subjects, that would mean that our estimated effect of feedback captures the effect of year unobservables that are not taken out by the year fixed effects. A possible explanation in that case, that would still facilitate our interpretation would be that students might react to feedback by studying more or less for the school instead of the national exams. But still, we would not be able to allay the concern that our estimated effect captures only the effect of feedback and not something else. In columns 3 and 4, we find that the coefficients are not statistically significant and there is no evidence that the provision of feedback affects students' performance in these subjects. What is important here is that students do not receive any social comparison information regarding the non-incentivized subjects neither in the feedback regime nor in the non-feedback regime. These findings are in line with our hypothesis that students change their effort choice and thus their next year performance due to receiving information about their relative performance, when feedback is provided.

We then run specification (2a) and in Figure 2.8 we plot the β_θ coefficients of the rank nationwide and the associated 95 % confidence interval. We observe that receiving information about someone's relative performance has a negative effect to students below the forty-fifth percentile and a positive effect to students above it. At the highest two deciles, the curve is slightly decreasing implying that there is a ceiling effect. In other words, there is some upper bound on how much improvement can feedback provision bring for the most able students. Thus, sitting similar exams prior to university admission high stake exams improves (decreases) the relative rank nationwide of the high (low) achieving students by up to 5 (8) percentiles. In Figure 2.9, we report γ_θ coefficients and the associated 95 % confidence interval. The estimated treatment effects on the rank within the school are very similar to the ones found before in Figure 2.8. This happens because the school sample that we use is a representative one in terms of many observed characteristics and so someone's rank nationwide might not differ a lot from his rank within his school. Figure 2.10 plots the treatment effect coefficients for the non-incentivized subjects that we use as the main non-treated subjects and we explained before. In line with Table 2.5, we find no evidence that the provision of feedback affects students' performance in these subjects.

The original score in each subject varies from 0 to 20. We standardise the twelfth grade scores in each year and school so that is has a zero mean and a standard deviation of one. The treatment effects line for each decile of prior performance is presented in Figure 2.11. There, the gain (loss) for students above (below) the 40th percentile is up to 0.15 (0.3) standard deviations.

2.5.2 Gender and Track

Next, we turn into the gender analysis. As literature on evaluating social programs has shown, individuals respond differently to the same policy (Heckman 2001). To test whether boys react differently than girls to the provision of feedback, we estimate the following regression:

$$Rank - nationwide_{ic} = \delta + \beta Feedback_c * Female_i + \kappa Feedback_c$$
$$+\lambda Female_i + \alpha X_i + \mu_t + \epsilon_{ic} \text{ (3a)}$$

$$Rank - school_{isc} = \delta + \beta Feedback_c * Female_i + \kappa Feedback_c$$
$$+\lambda Female_i + \alpha X_i + \mu_t + \epsilon_{isc} \text{ (3b)}$$

where X_i includes the tenth grade GPA performance, a dummy for early enrolment in school and dummies for the speciality chosen in the twelfth grade. OLS results are shown in Table 2.7. Although girls outperform boys, girls end up in a lower later rank on average when feedback is provided. This is the case when we consider both; the rank nationwide and the rank within their school. [24] Running specification (2b) [25] for boys and girls separately, produces Figure 2.13 that presents the treatment lines for boys (on the left) and girls (on the right).

For both genders, the effect of feedback is positive for high achieving students and negative for low achieving students. We make two important points here: First, in Table 2.7 the average effect of feedback on boys' last year rank is positive and on girls' is negative as shown by the horizontal line which is generated by a regression across all deciles (Figure 2.13). Second, the effects of feedback are more pronounced for women. As indicated by the steeper treatment line, girls exhibit greater sensitivity to knowing how well they do compared to their school or cohort peers.

Our evidence are consistent with the literature supporting differential gender effect to feedback with females responding more to additional information. McCarty 1986 in an experimental context, shows that women may react differently than men in the absence of feedback information because of different levels of self-confidence. Using an experimental context too, Franz et al. 2009 argue that women never have the same level of self-confidence as men because women expect less of themselves than men do.

Then, we disaggregate the analysis at the Track level. There are three Tracks: Classics, Exact Science and Information Technology and students have to take four exams within each track [26]. In Figure 2.16, we run specification (2a) separately for each track. The least average effect is observed for students in the Science Track whereas the treatment curve is steeper for students in the Information Technology Track rather than in Classics.

[24] In Table 2.7, if we include school fixed effects in specifications (3) and (4), we account for heterogeneity across schools and the coefficient estimates become the same as in columns(1) and (2).

[25] (2a) gives almost identical results as (2b) for both genders.

[26] In Classics they take national exams in: Ancient Greek, Latin, Literature and History. In Science the examined subjects are: Mathematics, Physics, Chemistry and Biology and in Information Technology: Computers, Mathematics, Physics and Business Administration.

2.5.3 Long Term Outcomes

In this section we examine the effect of feedback provision on students' long term outcomes. We have already motivated the discussion regarding the reasons a student would choose to resit the national exams for university admission. We use binary response models to examine if the provision of feedback affects the decision to retake the exam. In Table 2.10, we observe that a significant percentage out of the cohort population repeats the exam one year after graduation from senior high school[27].

We define as "misplacement" the difference between the tenth grade rank within the school each student gets and the rank nationwide in the twelfth grade. Thus, the misplacement variable is bounded between minus one and one. Students with larger differences between the tenth and the twelfth grade ranks would have a large change in their relative performance. The misplacement variable takes the value zero for students where their twelfth grade rank happens to correspond exactly to the tenth grade rank. But it can also take positive (negative) values if the student achieves a better (worse) relative performance in the tenth grade relative to the twelfth.

In order to examine if feedback provision affects someone's decision to retake the national exams through the misplacement effect we run the following specification:

$$Retake_{i,t+1,s,d} = a + X'_{itsd}\gamma + \delta Misplacement_{itsd}Feedback_t + \beta Feedback_t$$

$$+\omega Misplacement_{itsd} + \zeta Z_{td} + \xi_s + \omega_t + \epsilon_{itsd}$$

The decision to retake the national exam one year after graduation depends also on the opportunity cost of the student. Thus, we control for the unemployment rate in each year t and district d of student's residence.

Using Linear Probability (LPM), Probit and Logit models we find that when feedback is provided, students with higher misplacement are more likely to repeat the national exams one year after graduation. In Table 2.11, we interact dummies that capture the magnitude of misplacement with the feedback dummy and we observe that students in the top misplacement quintile (5) are more likely to resit the national exams when feedback is provided. The Top Misplacement Quintile (5) is the most positive one and contains students who get a better rank in the tenth grade compared to the twelfth. In the feedback years, these are the low achieving

[27] The number of students retaking the exam is calculated using the Ministry of Education dataset. The data about the labour force capacity are collected from the National Statistical Authority.

students. In other words, low (high) achieving students are more (less) likely to resit the national exams when feedback is provided.

Having a particular placement in university admission affects the employment and earnings prospects of an individual. We examine if feedback influences the matching of students to University Departments. We first rank all programs [28] according to their average cut-off over the seven years. Each program's cut-off expresses the society's valuation for this particular university department. Highly demanded programs exhibit high cut-offs. Students apply to programs based on preferences, social status and expected earnings. There are 659 programs in total. We estimate the effect of feedback on the difference in the popularity position and rank of the program admitted conditional on tenth grade performance. Figure 2.17 presents the treatment effect line for the popularity position (on the left) and rank of the program (on the right) admitted. The provision of feedback has a positive effect on the popularity position and rank of the program admitted in the upper half of the prior performance distribution and negative effect on the low half. In particular, high achieving students move up the University popularity ladder by 30 positions which is 0.15 of a standard deviation. Different placements in university admission induce different gains related to the returns to college.

We then use the 2003 Labour Force Survey to map each college major into the most related occupation and then into the expected annual earnings after graduation (in Euro).[29] In Figure 2.18, we present the effect of feedback on the expected annual earnings, conditional on the tenth grade performance. For students above the 50th percentile, there is an increase in their annual expected earnings by 250 Euros per year, which is equivalent to 0.17 standard deviation. For students below the 50th percentile, the decline in their expected annual earnings corresponds to 0.20 standard deviations.

2.5.4 Social Mobility

In this section, we examine if the provision of feedback changes the relationship between parental income (proxied by neighborhood income) and post-secondary opportunities (indicated by the program the student enrols in). A priori, we might expect that students coming from more advantaged neighborhoods would have better chances of embarking on a better and more-selective program with higher expected returns than students coming from less-advantaged neighborhoods. Could the provision of feedback affect this flow of students from high-income families to

[28] By program we mean each combination of University Department.

[29] Mean: 12,758 with 1,473 standard deviation.

high-expected-income programs? Providing relative performance information might have a different effect on students whose parents have varying levels of income; the difference in the role feedback plays may be related to other family resources (financial support or social networks) or students from different income backgrounds might value the ranking information differently.

To investigate whether feedback has a differential effect on students from different income backgrounds, we create quintiles based on the neighborhood income and the selectiveness of the program admitted to. In Table 2.16 we report for each quintile of neighborhood income, the percentage of students who enroll into each quintile of programs by selectiveness, in the feedback and the non-feedback regime. We then calculate the difference between the feedback and the non-feedback percentage. In the last row of Table 2.16, we vertically add the percentages of students who enroll in any program for each quintile of neighborhood income, to find the total difference of enrolled students between the feedback and the non-feedback period. In the last column of Table 2.16, we horizontally add the percentages of students who enroll in each quintile of programs. We do that to examine if feedback provision affects the total percentage of students who enroll in higher education. We find that 2.2 % more students (83.7% Vs 81.5%) enroll in a program in the feedback regime.

In Table 2.16, we find descriptive evidence that more students coming from the lowest-income neighborhoods (Quintile1) enroll in any program when feedback is provided (2.2% more students). A possible explanation is that low achieving students discover that if they do not exert more effort they will not be admitted to any program in tertiary education. Or they might discover that they are not worse than the low achieving students from high income neighborhoods and that they still have a chance to enroll in university. So, they might decide to exert more effort. This may show that feedback benefits students from low-income neighborhoods by reducing social inequalities and possibly future income inequalities. On the other hand, high achieving students from low-income neighborhoods discover in the eleventh grade that they are highly ranked on a national scale and they might react by exerting more effort.

We also find descriptive evidence in Table 2.16, that feedback provision alters the parental income (proxied by the neighborhood income) composition of students who are admitted into the top-ranked programs (Quintile 5). More students from low income neighborhoods are admitted to the most-selective programs that provide students with the highest expected earnings after graduation (such as engineering and law), when feedback information is provided (2.9% Vs 2.6 %). This implies that providing relative performance information encourages social elevation and improves economic opportunity for these students.

It is crucial from a policy perspective to understand if providing feedback is beneficial for the society as a whole. On one hand, high performing students are usually the ones responsible for innovation and technological breakthroughs. The technological diffusion is beneficial for the society as a whole, because technological innovation is one of the driving forces behind a country's economic prosperity and productivity advance (Nickell and Van Reenen [2001]). On the other hand, our study shows that providing relative performance information improves the performance of high-achieving students whereas low-achieving students perform even worse. This widening of the performance gap caused by feedback may be translated into a wage gap later. We find evidence to this direction using students expected wages. This might be detrimental especially for low achieving students. An economist may be fond of the efficiency achieved though information provision as high achieving students end up higher in the society and the spillover effects of the technological advances to the whole society. Nevertheless, at the end of the day its up to the society to decide whether efficiency can be traded for equality.

Additionally, our descriptive statistics evidence show that providing the relative performance information may encourage students from low income families to enroll in university and especially to more selective programs. From this perspective, providing the relative performance information encourages social elevation for students coming from low income neighborhoods. Thus, feedback decreases the performance or income inequality between students coming from low and high income neighborhoods.

2.5.5 Positive Vs Negative Surprise

In this section, we examine whether students respond to the specific type of feedback that they get. Students might not only compare themselves with their class or school or cohort-mates but they may also compare their own relative performance in different periods in time. Exploiting within school variation in the 134 senior high schools and the fact that we know the whole distribution of scores, we restrict this part of the analysis into the feedback years. If a student receives information that he is in a higher (lower) decile in the eleventh grade than in the tenth grade, then our student receives a positive (negative) shock, that can be translated into a "positive (negative) surprise". Intuitively, students who receive a positive (negative) surprise in the eleventh grade might increase (decrease) their expectations of themselves and exert more (less) effort in the twelfth grade. In order to examine potential effects coming from the surprise they experience in the eleventh grade, we graph the effect on the twelfth grade rank for each combination of percentile ranks in the tenth grade and eleventh grade. That is shown by the hitplot in Figure 2.19.

The horizontal axis represent the eleventh grade percentile rank of students and the vertical axis represent the tenth grade percentile rank. Different colours express different magnitudes of the treatment effects on the twelfth grade rank. The diagonal starting from zero towards the right upper edge of the box, represents the case of "no value feedback" or in other words those students with their eleventh grade percentile rank equal to the tenth grade percentile rank. The treatment effect is positive (negative) for most students experiencing a positive (negative) surprise.

A concern here is that students might not be aware of their tenth grade percentile rank, especially if they attend a school with more than one classes. However, the analysis here is done for deciles of performance and not for percentiles, allowing students to have priors that do not accurately express their exact tenth grade rank.

2.5.6 Alternative Mechanism: School quality revelation

An alternative mechanism could be that students use the information obtained by the publication of their scores in such a way that they realise the quality of their senior high school [30] Students who take the eleventh grade national exams suddenly realise their school rank and their national rank and the comparison of the two ranks reveals information about the quality of the school. If a student realises that his national rank is greater than the school rank then his school is of good quality. The opposite if the national rank is lower than the school rank. The realisation of the school quality in the eleventh grade might affect students' choice of effort in the twelfth grade. Thus, we exploit the across schools variation in their quality to identify the effect of feedback on students' rank nationwide.

In Figure 2.20, we produce the treatment lines separately for students who realise that the school they attend is worse (on the left) and better (on the right) than the average quality school. We find the effect of providing the relative performance information on these students' standardised national exam score. [31] The average effect for students who realise that they attend a worse than average quality school is negative whereas it is positive for those who realise that they attend a better than average quality school.

[30] We measure school quality based on the schools' average national exam performance in the twelfth grade from 2003 to 2009. Then we construct a rank measure for school quality that varies from zero to one. The average quality of the schools in our sample is 0.52 (sd 0.21, Min 0.018 and Max 0.985) which means that our school sample is of a representative quality.

[31] Standardised within each year with zero mean and a standard deviation of one.

Starting with the bottom of the prior performance distribution, we observe that low achieving students in good schools do better that those in lower quality schools. Surprisingly, there is a huge increase in the national rank for the top students in the worse schools and this increase even offsets the increase in the national rank of the top students in good schools. We acknowledge two possible explanations here: First, the better students in the worse schools take the eleventh grade national exams, they receive feedback, they realise that they are actually exceptional in a national scale and thus decide to exert more effort in the next time period. So feedback acts as a motivation boost for these students.

Second, the realisation of their national rank act as a rude awakening for these students who might initially have a wrong perception about the national competition and about their school's quality. These students might be the top students in their class or school but they now learn that they are left behind. In the next time period, they exert more effort in order to catch up with the national competition.

The fact that the two curves do not follow the same pattern enhances the argument that the results are not driven by experience or practice resulted from sitting the eleventh grade national exam. If students realise the quality of the high school through the eleventh grade national exams, then would all receive the same information and they would not react so differently.

2.5.7 Alternative Mechanism: Practice

It could be argued that students can accurately place themselves within their class, even if they are not explicitly informed about their within-class rank. This is likely to occur due to repeated interactions among classmates throughout high school. However, here students receive new information about their relative performance in reference groups broader than the class. Consider the within school rank: students receive information about how well they did within their school. In Figure 2.21, we report the treatment lines for students in schools of different capacity in the eleventh grade. We make four broad categorisations. First, we consider schools with only one class where it is likely that students already know their relative standing and the social comparison information has no extra value (Panel A). Nevertheless, in a school with only two classes students might know their relative performance in their class but not in the whole school. Thus, we see that there is a small positive feedback effect on students who are above the 40th percentile and a small negative effect on those below it (Panel B). Additionally, the treatment lines become steeper when we consider schools with three classes (Panel C). In this case, the information is much broader than that which students can collect from interaction with their classmates. This is even more pronounced when we look at students in schools with more than three classes (Panel D). Summary statistics about the capacity of schools in our sample are presented in Table 6. Figure 2.21 shows that the effect of feedback depends on whether the additional information is actually informative about someone's relative performance.

That could allay the concern that the eleventh grade national exam might provide students with experience or training instead of information about their relative performance. School exams in the eleventh grade have the same format as national exams in the eleventh grade and the past papers are available in both regimes. Students practice on past questions and are aware of the structure and the types of questions in both cases. The structure of the exam paper is the same from one year to another for all years included in our sample. If students were experienced from sitting the eleventh grade national exams, then the experience or training effect would not vary by the size of the school. In other words, if that was the mechanism then students in small schools would have no reason to react differently than students in regular schools.

2.5.8 Alternative Mechanism : Parental investment

Another possible mechanism is that parents decide to invest more or less in students based on the eleventh grade results. Parents may start devoting more time helping the child with the homework or they may invest in external support (such as supplementary material/books, private tutors etc). It is true that there might be variation in family income within the school and the neighborhood income represents the average income in each region. We observe considerable differences in neighborhood income.[32]

In Figure 2.22, we draw the treatment lines for the bottom and the top quintile of neighbordhood income. We run specification (2b) separately for the top and the bottom quintiles of neighborhood income. This may not fully reflect the family income but we examine the effects of feedback across regions by average reported income. A wealthy family may have the financial resources to invest in the child and thus the student may improve his performance in the subsequent year exam. On the other hand, families from low income neighborhoods may not be able to pay enough to further support the student. In Figure 2.22 (right panel), we observe that disclosing rankings increases the average subsequent national rank for students coming from the highest-income neighborhoods. The average effect on the subsequent national rank for students from the lowest-income neighborhoods

[32] mean: 23,517 standard deviation: 8,609 min.:13,005 max.: 66,521

is negative. In high-income neighborhoods the positive effects of feedback hold for students above the 40th percentile while only students above the 60th percentile from low income neighborhoods benefit from feedback. If parental investment was the only driver of the findings, then we would expect students from highest-income neighborhoods to improve at all parts of the prior-performance distribution. That implies that there might be, to some extent, differential parental investment in students by family income (proxied by neighborhood income) but that cannot fully generate our results.

2.6 Threats to Identification

2.6.1 Attrition

In our attempt to evaluate the impact of feedback on different performance groups, the problem of attrition cannot be ignored. If attrition is random and affects different performance groups in a similar way in both regimes, then the estimates remain unbiased. Differential attrition here could arise because students from the lowest percentiles are more likely to drop out from school in comparison to students from the highest percentiles when they realise their relative ability performance. If that is stable over time, it will not affect our feedback estimates. What could bias our estimates, is if differential attrition follows the abolition of feedback. [33] In Table 2.11 we report the drop out rate between tenth and eleventh grade as well the drop out rate as eleventh and twelfth grades. We also report the percentage of students who transfer to a school in the eleventh (column 2) and twelfth grade (column 4). We observe that there are no significant changes in the percentages of students who dropped out or transfer to another school before or after feedback was abolished.

Exploiting within school variation, we use the following specification to check for differential that changes with feedback:

$$Drop-out_{12-10isc} = \alpha + \beta_{quintile} Feedback_c * Quintiles_{10isc} + \lambda_{quintile} Quintiles_{10isc}$$

$$+\psi Feedback_c + X'\gamma + \theta_c + \varphi_s + \epsilon_{isc}$$

Table 2.13 reports OLS results. The drop out rate is larger for the lowest quintile than any other compared to the third quintile when feedback is provided. But more importantly, none of the coefficients of interest are statistically significant.

[33] The first affected cohort for which feedback is abolished is the cohort that was in the twelfth grade in 2006. Thus, this cohort was in the tenth grade in 2004. This is the first cohort that did not sit national exams in the eleventh grade.

This implies that there is differential attrition, but it does not vary with feedback policies.

2.6.2 Robustness checks

In this section, we construct a robustness exercise to complement our main analysis. One concern is that the change in the variation of performance over time might not be caused by the provision of feedback. In other words, we need to rule out the possibility that the better students become worse over time and the worse students become better over time for reasons different than the provision of feedback.

Exploiting the within school variation, we run specification (2b) but without pooling feedback and non-feedback years together. Instead, we just compare every pair of consecutive years in the sample. The only pair of years that we expect to find a differential response of cohorts is 2005-2006 (the year of the reform). For every other pair of years, we expect to find similar cohort behaviour. We present the placebo treatment lines in Figure 2.23. Panel A compares the cohort 2003 to the 2004, as if feedback was abolished in 2004. We find no evidence that other factors might affect students differently in other years. Panel C corresponds to the actual reform and we observe that the treatment effects are negative for all percentiles below the 50th percentile and positive above it. Regarding any policy anticipation effects, the reform was announced in around December of 2003-2004. We find very small treatment effects in Panel D, which is the first non-treated cohort. Students in the first non-treated cohort might observe how last year's peers of similar tenth grade performance did and use this information to adjust their effort. Again after 2007, the curve is almost flat throughout the ability distribution.

We conduct some other placebo exercises to verify that the effect does not depend on the numbers of subjects examined. In Figure 2.24, we draw the treatment lines for each subject separately. In Table 2.14, we calculate the final-year rank based on different subjects. In column (1) we find the effect of feedback on the final year rank that takes into account the Electives or Track subjects on the top of the core education subjects[34] and the results are very similar to those reported so far. In column (2) we take into account the effect of feedback on students' performance in Modern Greek which is a common subject in both regimes and takes a special weight in the calculation of the University admission grade. Notice, that in the non-feedback regime two subjects are examined nationally and three within the school. In column (3) we calculate the last year's rank based on the five subjects

[34]Students sit national exams in four Elective subjects. So the overall rank in calculated based on nine subjects.

in the feedback regime and the two subjects in the non-feedback regime. Results remain very similar. Treatment effect remain positive (negative) for the high (low) achieving students.

2.7 Conclusion

In this chapter, I examined the effects of providing relative performance information on students' short and long term outcomes. I exploit a large scale natural experiment that took place in Greece and thus we conducted a large scale primary data collection process. Using detailed data on students' performance throughout senior high school and school quality data, I examine the effects of receiving information about someone's national and school relative performance. It is human nature to make comparisons, which can affect students' beliefs about own ability and effort decisions. For students above (below) the 50th percentile, I found that feedback information has a positive (negative) effect on their subsequent performance, popularity of program admitted and expected annual earnings.

The actual reform in this chapter was the abolition of the relative performance information that was provided in the eleventh grade. I believe that the introduction of feedback would have the same effects as we find here. I think that these effects are driven by the relationship between effort and ability (the production function). In this setup, effort and ability seem to be complements in the production function. This means that high achieving students will get encouraged when their relative ability becomes known and they will exert more effort. Low achieving students will exert a lower effort when the information is provided due to discouragement effects. In a different setup, where ability and efforts are substitutes the effects would go to the opposite direction. For example, if students want to reach a specific performance threshold then high achieving students would have to smooth effort across the two grades or maybe even reduce the level of effort exerted in the twelfth grade. This entirely depends on the relationship between performance threshold and ability. If feedback is not the main mechanism then it is hard to predict what would happen if feedback was introduced (instead of abolished). However, in this study we find evidence that feedback is our main mechanism that drive the results.

Our estimates are at the lower end of those compared to the literature on improving school inputs. Nonetheless, all the interventions studied so far (improving teachers quality, reducing class size, enhancing the peer quality group) are significantly more costly than the provision of feedback. Releasing feedback information is a low cost instrument by which to raise students achievement in University admission exams.

I outlined two potential mechanisms in this study for why students would react to the provision of feedback. The first one supports that with feedback, students update their belief about their own relative ability and that determines the next period's effort choice, as explained in the theoretical model. Another possible mechanism is that students combine the country and school level information about their ranks that reveals new information about their school quality. Knowing the school quality might provide information to students about the level of the competition over restricted university places. I use these mechanisms to explain our results.

These findings have important policy implications both in relative and absolute terms.

First, the effects of feedback are positive on the high achieving students and negative on the low achieving students implying that policy makers need to be cautious depending on who they target. These effects concern students' next year performance but also long term outcomes. Feedback provision affects the matching with the university department students are admitted to and consequently their life term earnings. Secondly, girls are more sensitive to feedback and they respond more at both tails of the ability distribution. The relative nature of the above mentioned results restrict the broad implementation of feedback, but makes it very important in a competitive process. Our analysis highlights the importance of rank position on students' scholastic and labour market outcomes and we believe that the rank could be a new factor in the education production function.

The analysis moves on highlighting the absolute effects of feedback: First, high achieving students in worse schools gain a lot from feedback. Second, the consequence of no feedback is more resitting for high achieving students. This is an important loss of human capital for the society given that the most able students stay out of the university and/or the labour market. Third, I find evidence that feedback encourages students from low-income neighbourhoods. More precisely, more students from low-income regions gain admission to top University Departments when feedback is provided, indicating a potential future drop in income inequality.

2.8 List of Figures

Figure 2.1: Map of schools in the sample

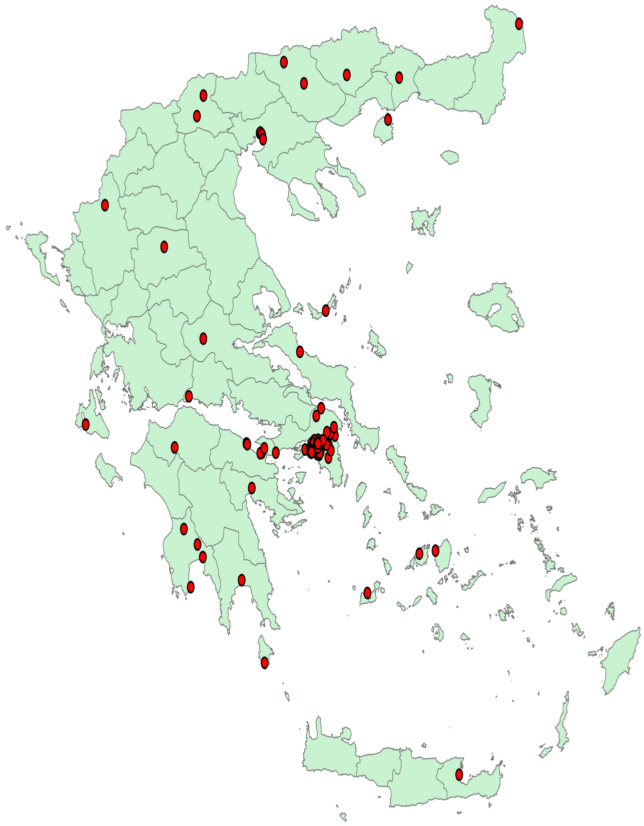


Figure 2.2: Timing

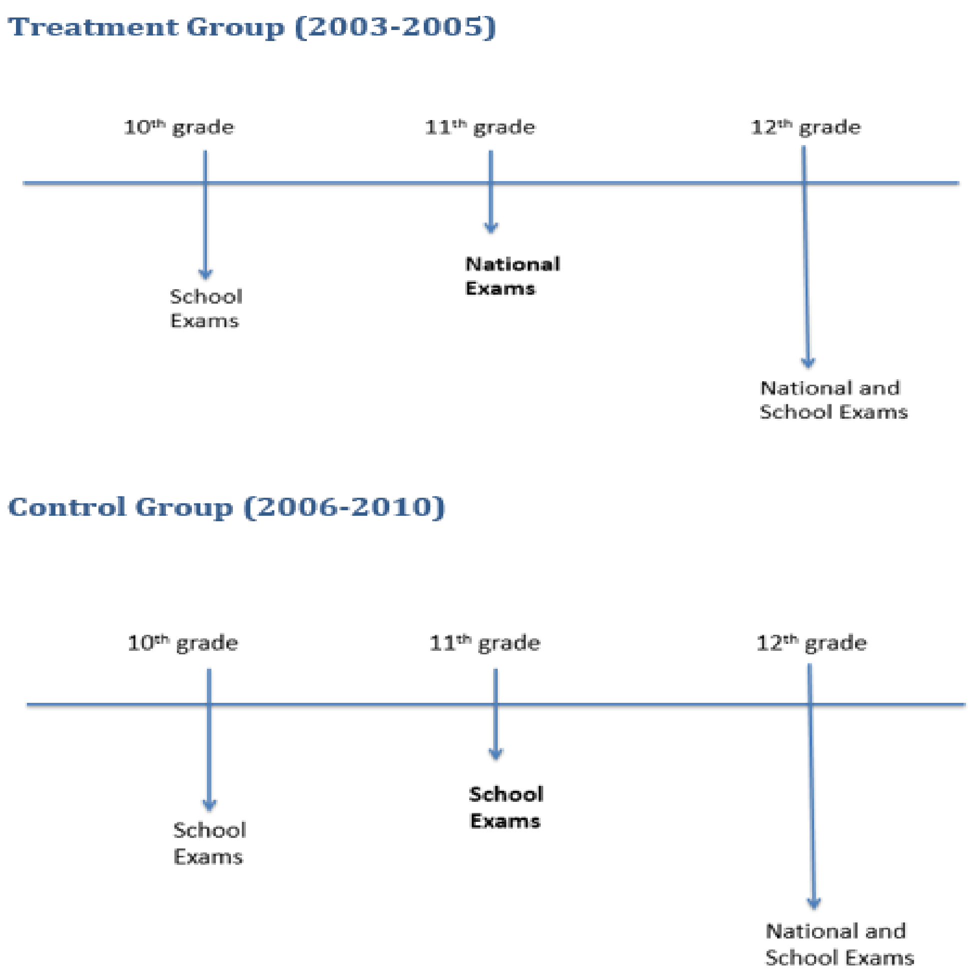


Figure 2.3: Announcement of school results

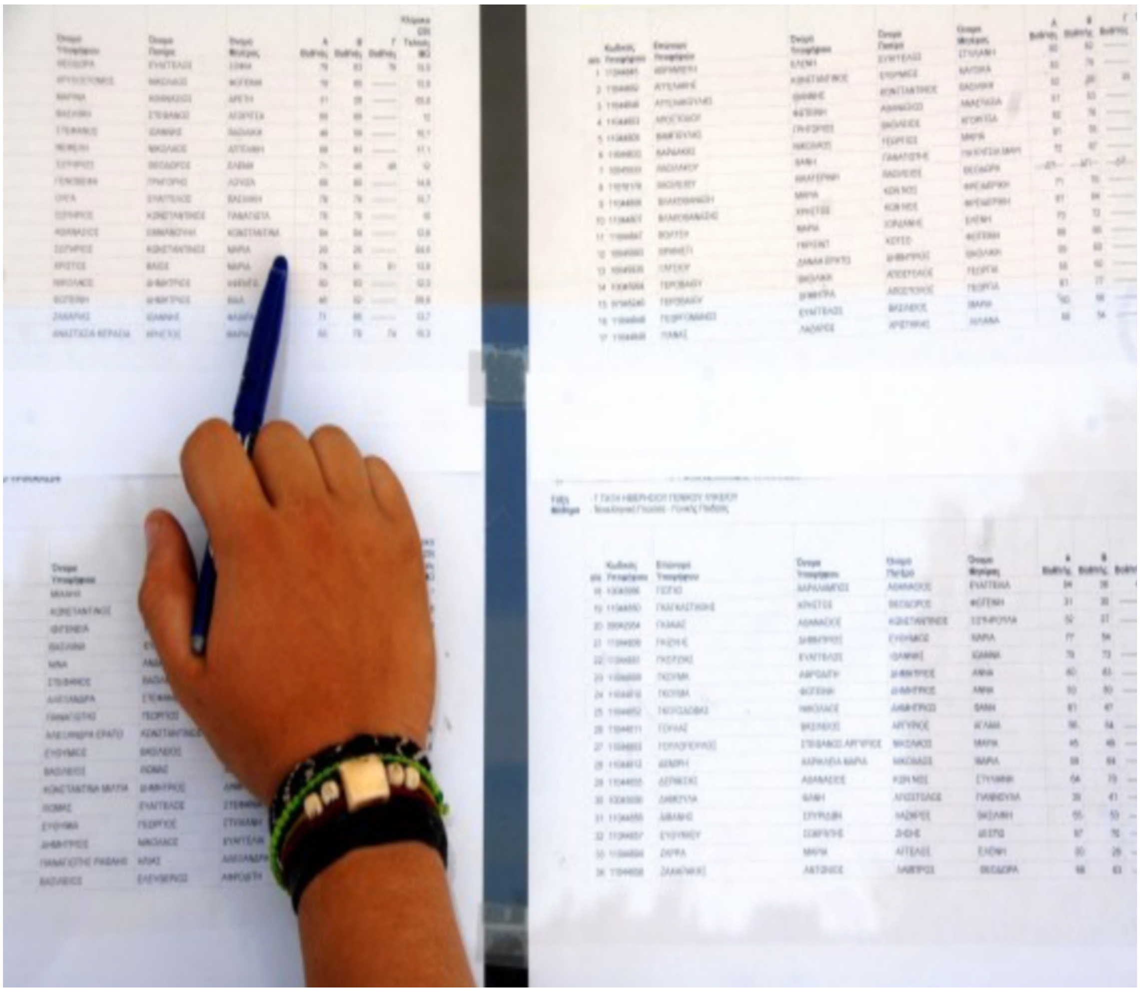
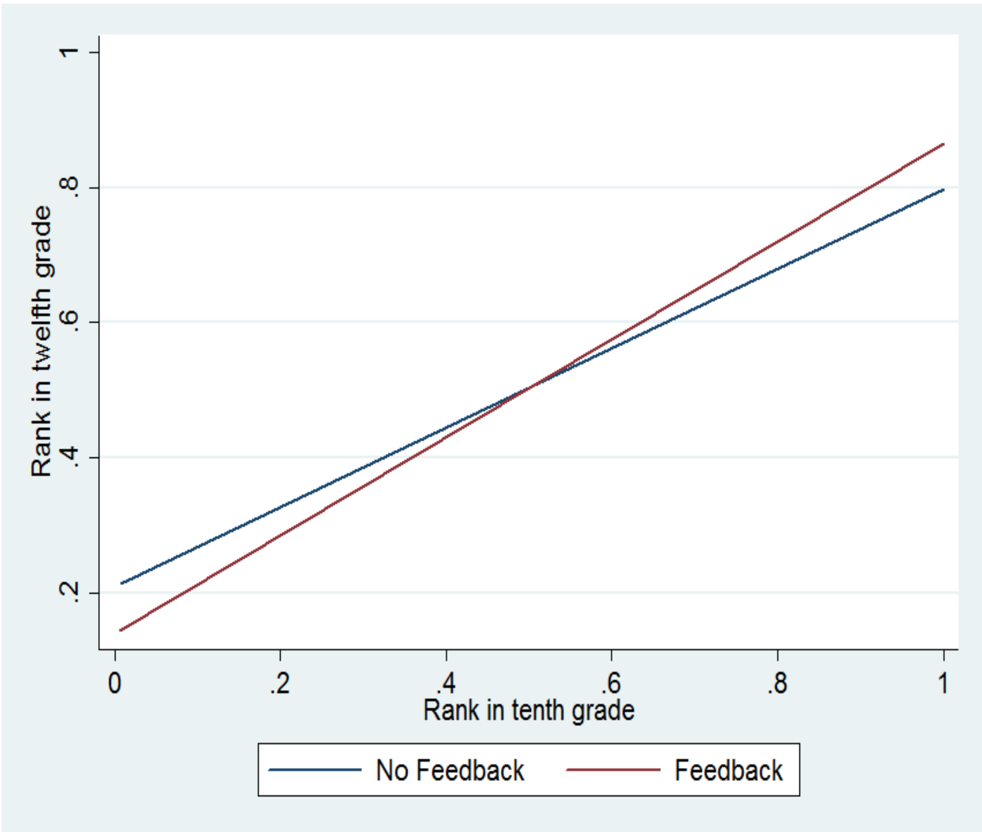


Figure 2.4: Announcement of school results-Zoom in

Α/Α	Κωδικός υποψηφίου	Επώνυμο	Όνομα	Όνομα πατρός	Α Βαθμός	Β Βαθμός	Γ Βαθμός	Τελικός Βαθμός
2	15030628	ΑΛΑΒΕΡΑ	ΜΑΡΙΑ ΕΛΕΝΗ	ΙΩΑΝΝΗΣ	84	79		16,3
4	15030629	ΒΑΡΚΑ	ΙΩΑΝΝΑ	ΘΕΟΔΟΣΙΟΣ	93	84		17,7
8	15030630	ΔΕΜΑΓΚΟΣ	ΚΩΝΣΤΑΝΤΙΝΟΣ	ΙΩΑΝΝΗΣ	27	38		6,5
14	15030631	ΕΞΑΡΧΟΠΟΥΛΟΥ	ΠΑΝΩΡΑΙΑ	ΔΗΜΗΤΡΙΟΣ	80	89		16,9
17	15030632	ΖΕΓΙΑΝΝΗ	ΜΑΡΙΑ ΠΑΝΑΓΙΩΤΑ	ΜΙΧΑΗΛ ΣΤΑΜΑΤΙΟΣ	51	45		9,6
19	15030633	ΚΑΠΠΑ	ΙΩΑΝΝΑ	ΠΕΤΡΟΣ	73	74		14,7
21	15030634	ΚΑΡΑΚΑΤΣΙΩΝΗ	ΒΑΣΙΛΙΚΗ	ΚΩΝΣΤΑΝΤΙΝΟΣ	33	22		5,5
22	15030635	ΚΑΡΠΑΘΑΚΗΣ	ΚΩΝΣΤΑΝΤΙΝΟΣ	ΙΩΑΝΝΗΣ	77	57	70	14,7
23	15030636	ΚΑΡΠΑΘΙΟΥ	ΝΟΜΙΚΗ ΗΛΙΑΝΑ	ΜΗΝΑΣ	40	34		7,4
24	15030637	ΚΑΤΑΚΑΛΕΑΣ	ΑΝΤΩΝΙΟΣ	ΠΑΝΤΕΛΗΣ	82	90		17,2
25	15030638	ΚΑΤΣΑΜΑΚΗ	ΧΑΡΑ ΣΕΒΑΣΤΙΑΝΑ	ΕΥΠΡΕΠΙΟΣ	52	48		10
29	15030639	ΚΟΥΜΠΟΓΙΑΝΝΗΣ	ΑΝΑΣΤΑΣΙΟΣ	ΑΘΑΝΑΣΙΟΣ	90	93		18,3
31	15030640	ΚΡΙΜΙΖΗ	ΧΡΥΣΟΒΑΛΑΝΤΟΥ ΔΟΜΝΙΚΗ	ΜΙΧΑΗΛ	28	24		5,2
32	15030641	ΚΥΔΩΝΑΚΗ	ΣΟΦΙΑ	ΑΝΤΩΝΙΟΣ	80	91		17,1
34	15030646	ΜΑΥΡΙΔΟΥ	ΑΙΚΑΤΕΡΙΝΗ	ΑΛΕΞΑΝΔΡΟΣ	14	15		2,9
43	15030648	ΝΤΙΝΩΡΗ	ΕΛΕΥΘΕΡΙΑ ΣΑΒΒΟΥΛΑ	ΠΑΝΑΓΙΩΤΗΣ	91	95		18,6
45	15030649	ΟΙΚΟΝΟΜΟΥ	ΑΝΝΑ ΦΙΛΙΑ	ΝΙΚΟΛΑΟΣ	96	94		19
51	15030650	ΠΑΡΒΕΡΗ	ΣΕΒΑΣΤΗ ΕΥΑΓΓΕΛΙΑ	ΑΝΑΣΤΑΣΙΟΣ	87	88		17,5
57	15030652	ΣΜΑΡΑΓΔΑΚΗ	ΕΥΑΓΓΕΛΙΑ	ΓΕΩΡΓΙΟΣ	100	99		19,9
58	15030656	ΣΩΤΗΡΙΔΗΣ	ΔΗΜΗΤΡΙΟΣ	ΓΕΡΑΝΤΙΟΣ	77	77		15,4
60	15030659	ΤΡΙΑΝΤΑΦΥΛΛΟΥ	ΔΗΜΗΤΡΑ	ΘΕΜΕΛΗΣ	67	53	68	13,5

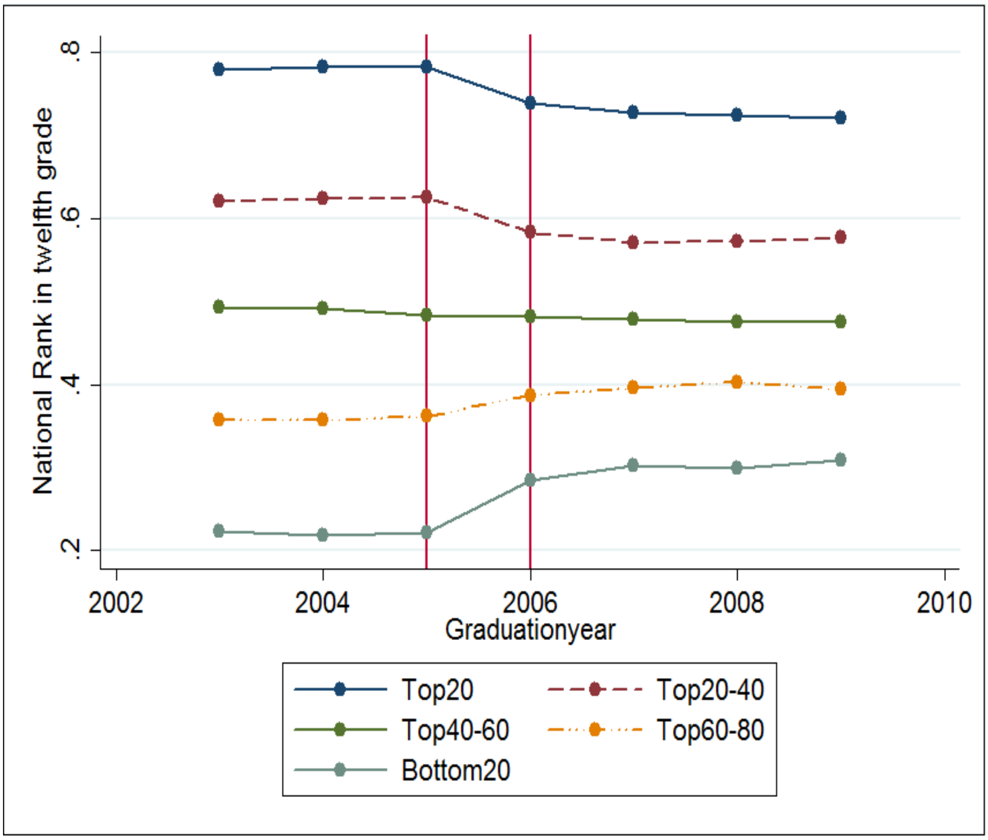
Note: This is the format of the publicly announced results. In particular, publicly announced are the following: a random code, a unique code, a student's surname, a student's first name, the father's name, the score given by the first examiner, the score given by the second examiner, the score given by the third examiner-if necessary- and the final score. The final score is the average of the score given by the first and second examiners. Students are sorted by alphabetical order. If the difference between the score given by the first and the second examiner is greater than or equal to 13, then a third examiner is required. The scores given by the first, second and third-if necessary- examiners are from 0 to 100 while the final score ranges from 0 to 20. If a third examiner is required, then the final score is the average between the two highest scores given by any examiners.

Figure 2.5: Fitted values



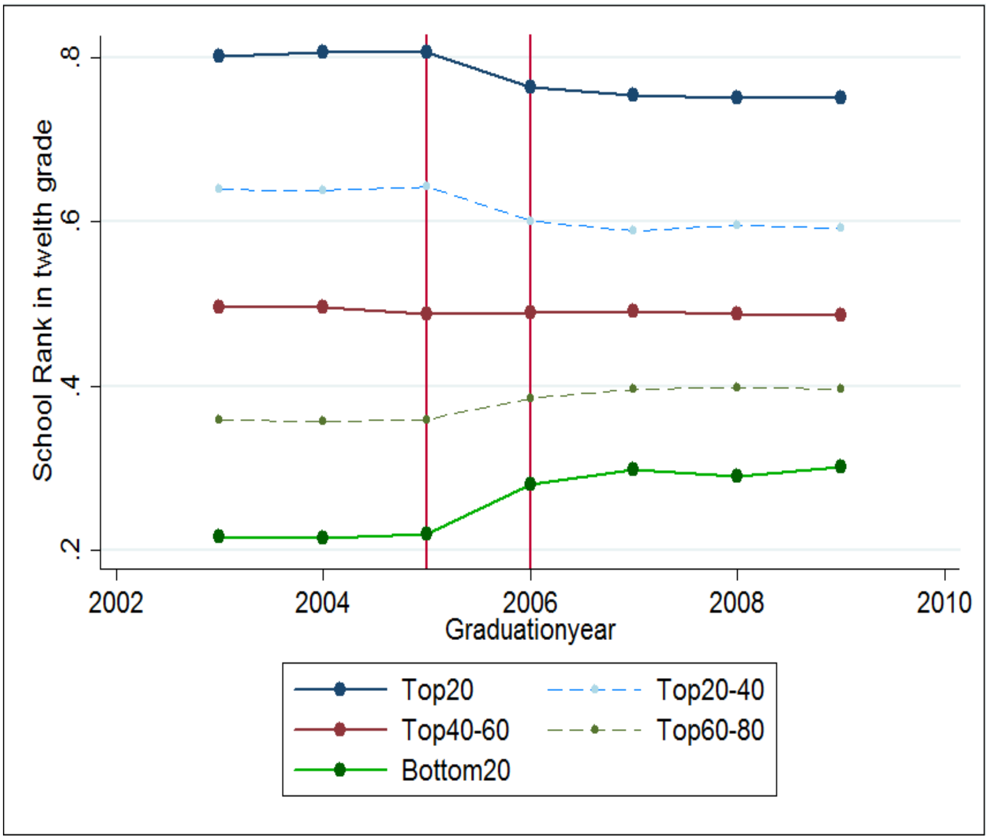
Note: Feedback regime refers to years from 2003 to 2005 while non-feedback regime refers to years 2006-2009. The rank in the twelfth grade is calculated based on the five incentivized subjects while the rank in the tenth grade is based on the tenth grade GPA.

Figure 2.6: Time trends for twelfth grade rank nationwide



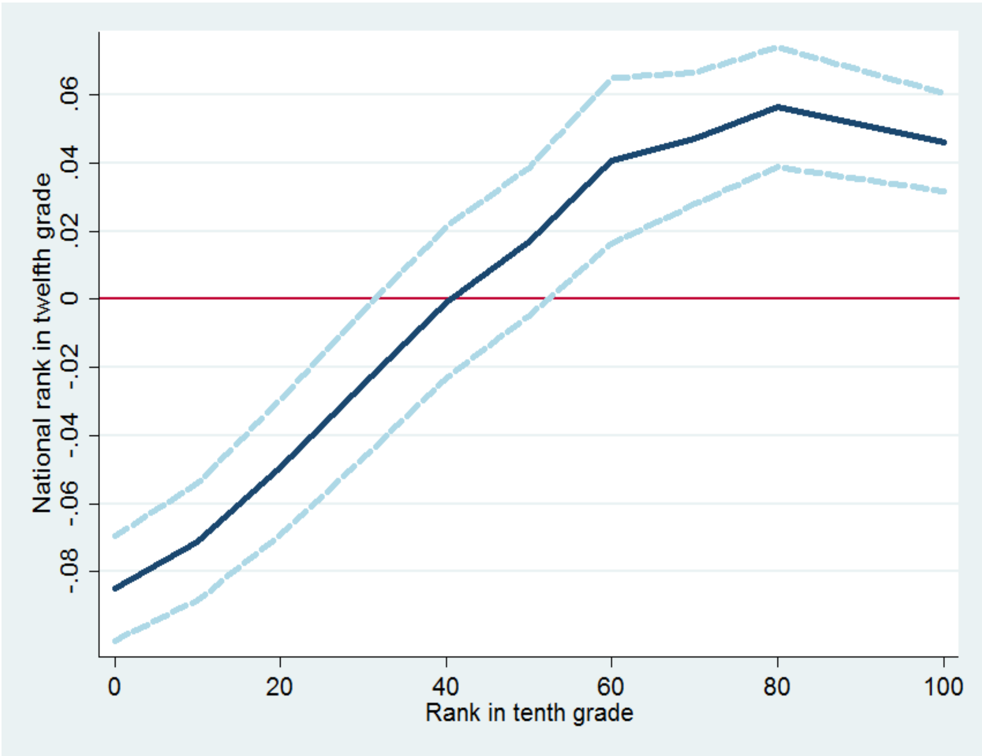
Note: Feedback provision for cohorts 2003-2005.The 2006 cohort is the first one for which feedback is abolished. Outcome variable: The national rank in twelfth grade. The trends correspond to different performance groups based on the tenth grade performance.

Figure 2.7: Time trends for twelfth grade rank within the school



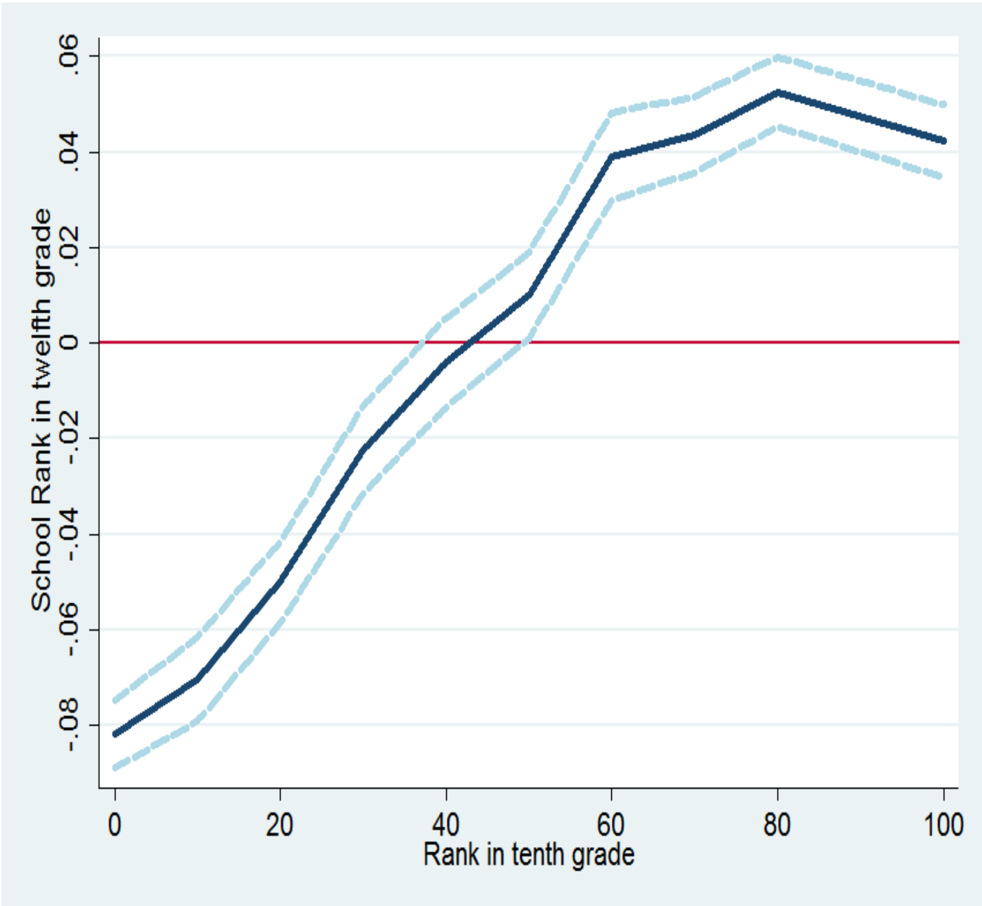
Note: Feedback provision for cohorts 2003-2005.The 2006 cohort is the first one for which feedback is abolished. Outcome variable: The rank in twelfth grade within the school. The trends correspond to different performance groups based on the tenth grade performance.

Figure 2.8: Treatment effects on the rank nationwide conditional on prior performance



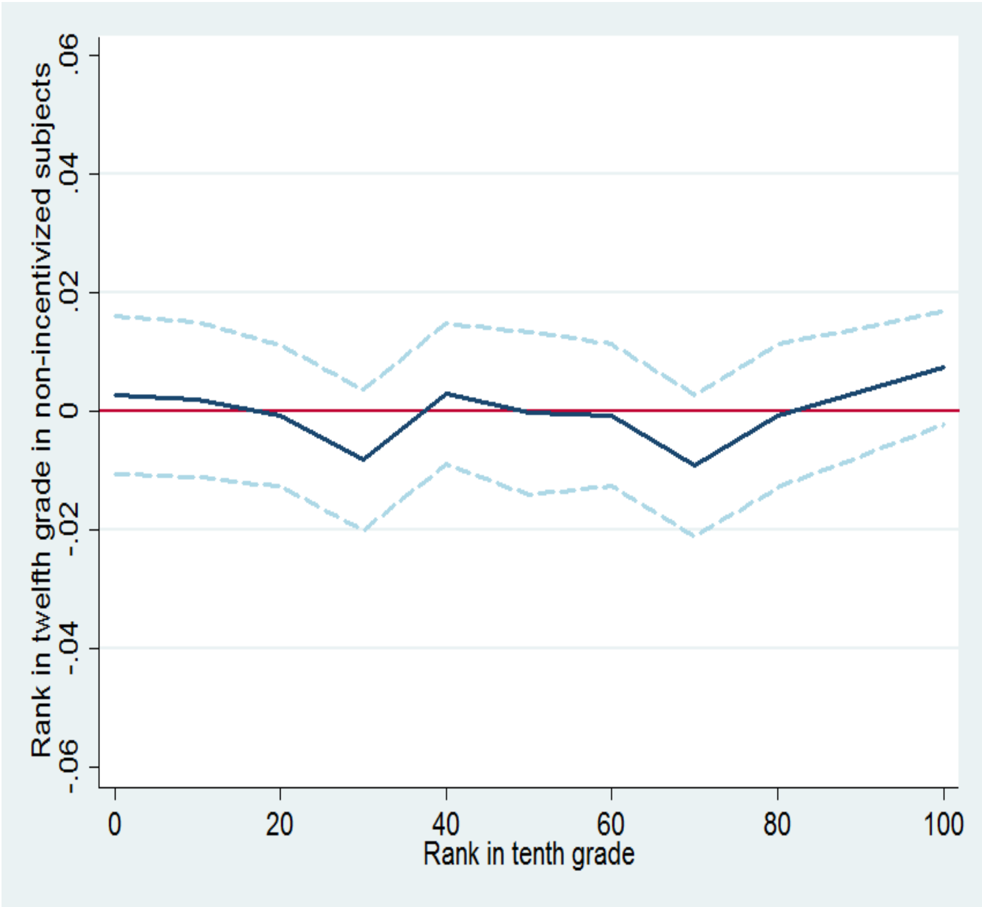
Note: The estimated effect of feedback on the national rank in the twelfth grade at each decile of students' GPA performance in the tenth grade and the associated 95 % confidence interval. The national rank is calculated based on the five core educational subjects (incentivized). The regressions are conditional on the students' characteristics: gender, age, a dummy that takes the value of one if the student is early enrolled in school and dummies for the track each student chooses in the twelfth grade. Standard errors are clustered at the school level.

Figure 2.9: Treatment effects on the rank within the school in incentivized subjects conditional on prior performance



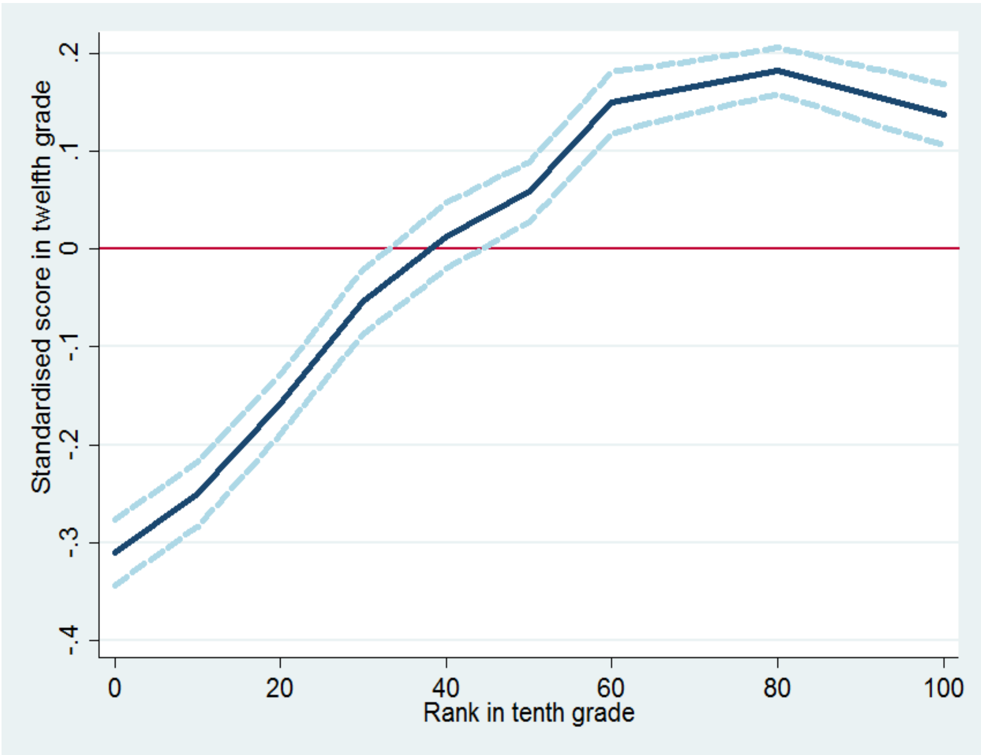
Note: The estimated effect of feedback on the school rank in the twelfth grade at each decile of students' GPA performance in the tenth grade and the associated 95 % confidence interval. The school rank is calculated based on the five core educational subjects that students take in the twelfth grade and determine the University admission grade (incentivized subjects). The regressions are conditional on the students' characteristics: gender, age, a dummy that takes the value of one if the student is early enrolled in school and dummies for the track each student chooses in the twelfth grade. Standard errors are clustered at the school level.

Figure 2.10: Treatment effects on the rank within the school in non-incentivized subjects conditional on prior performance



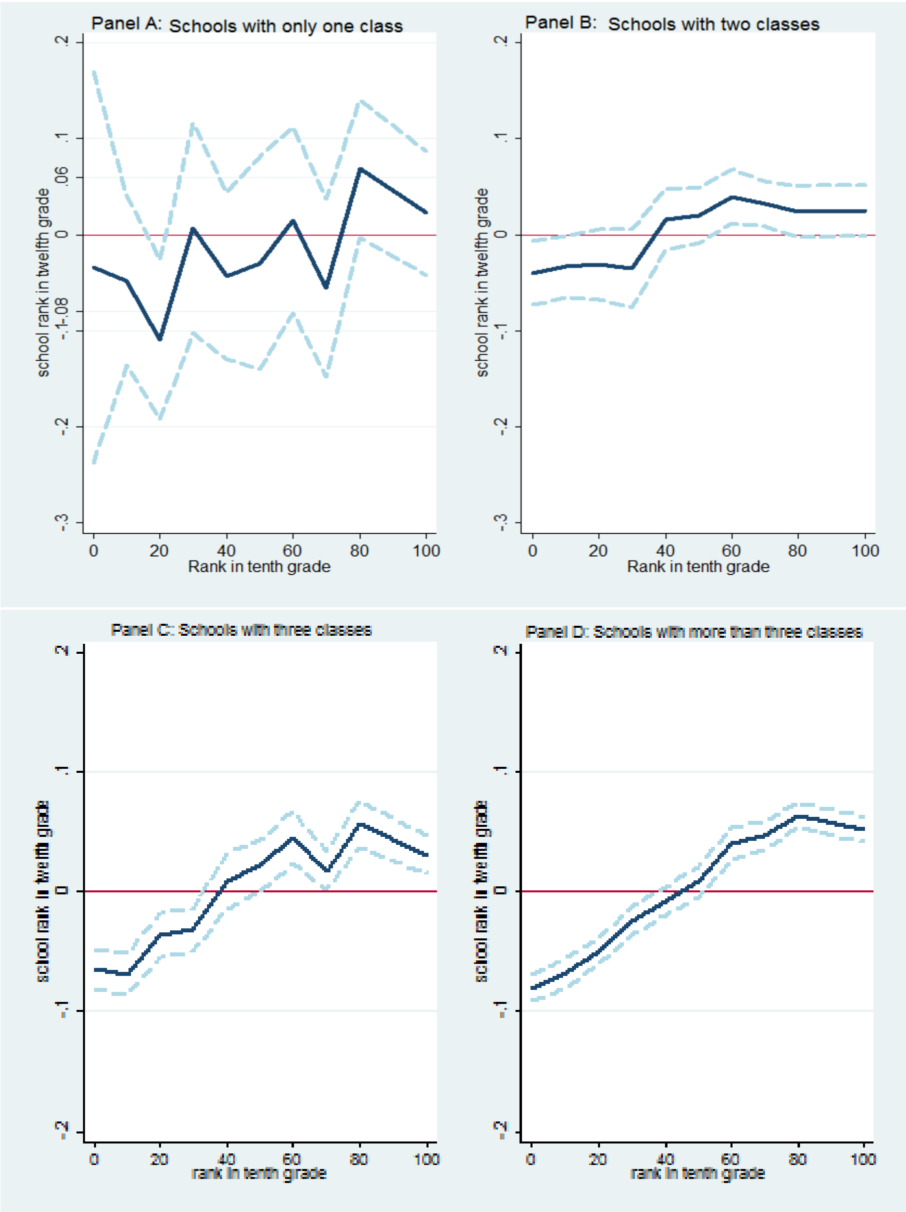
Note: The estimated effect of feedback on the school rank in the school exams at each decile of students' GPA performance in the tenth grade and the associated 95 % confidence interval. The school rank in the school exams is calculated based on the three non-incentivised subjects that all students take in the twelfth grade and these subjects are not taken into account in the calculation of the University admission grade. Students never receive social comparison information in these subjects. The regressions are conditional on the students' characteristics: gender, age, a dummy that takes the value of one if the student is early enrolled in school and dummies for the track each student chooses in the twelfth grade. Standard errors are clustered at the school level.

Figure 2.11: Treatment effects on the standardised score conditional on prior performance



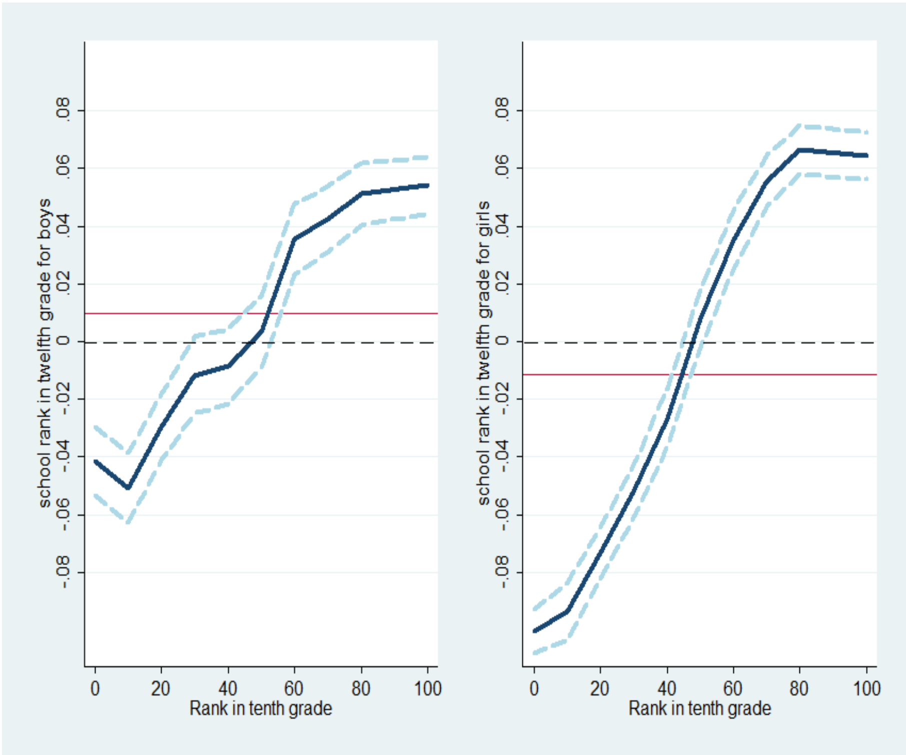
Note: The estimated effect of feedback on the standardised score in the twelfth grade at each decile of students' GPA performance in the tenth grade and the associated 95 % confidence interval. The standardised score is calculated based on the five core educational subjects (incentivized). The standardised score has a mean of zero and a standard deviation of one in each year. The regressions are conditional on the students' characteristics: gender, age, a dummy that takes the value of one if the student is early enrolled in school and dummies for the track each student chooses in the twelfth grade. Standard errors are clustered at the school level.

Figure 2.12: Treatment effects on the rank within the school conditional on prior performance for schools of different capacity



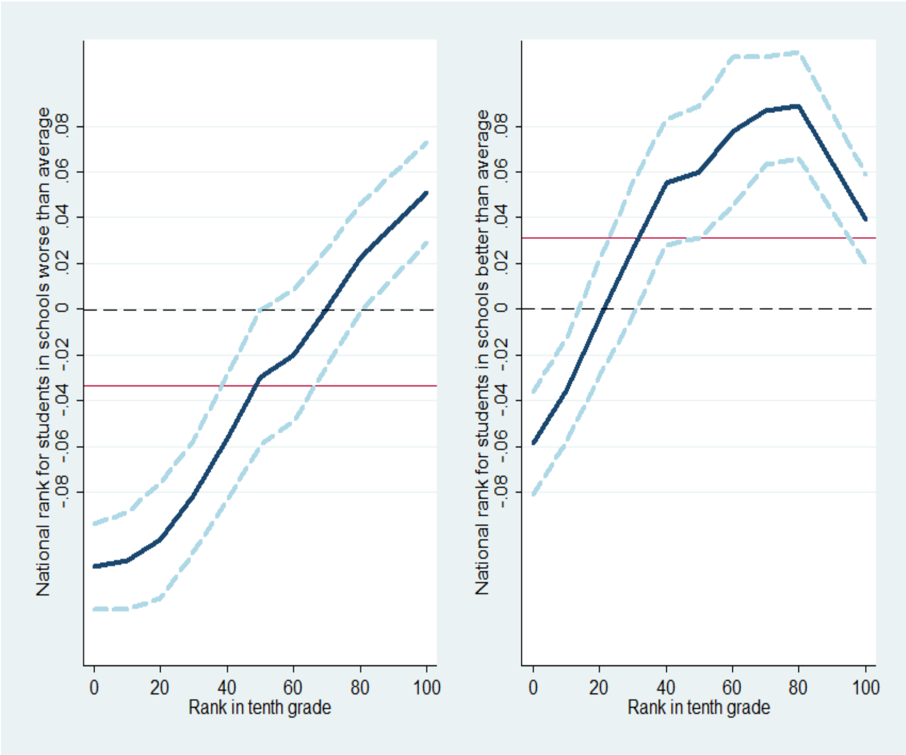
Note: The estimated effect of feedback on the school rank in the twelfth grade by capacity of school at each decile of students' GPA performance in the tenth grade and the associated 95 % confidence interval. The regressions are conditional on the students' characteristics: gender, age, a dummy that takes the value of one if the student is early enrolled in school and dummies for the track each student chooses in the twelfth grade. Standard errors are clustered at the school level.

Figure 2.13: Treatment effects on the rank within the school by gender conditional on prior performance



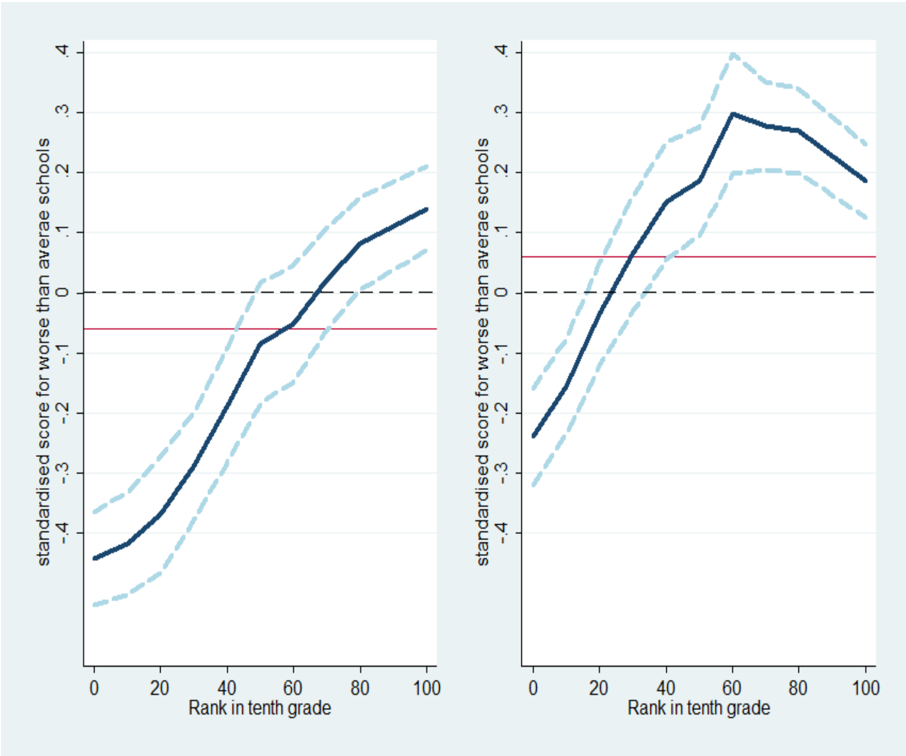
Note: The estimated effect of feedback on the school rank in the twelfth grade by gender at each decile of students' GPA performance in the tenth grade and the associated 95 % confidence interval. Males are depicted on the left and Females on the right. The school rank is calculated based on the five core educational subjects (incentivised). The regressions are conditional on the students' characteristics: gender, age, a dummy that takes the value of one if the student is early enrolled in school and dummies for the track each student chooses in the twelfth grade. Standard errors are clustered at the school level.

Figure 2.14: Treatment effects on the rank nationwide by school quality conditional on prior performance



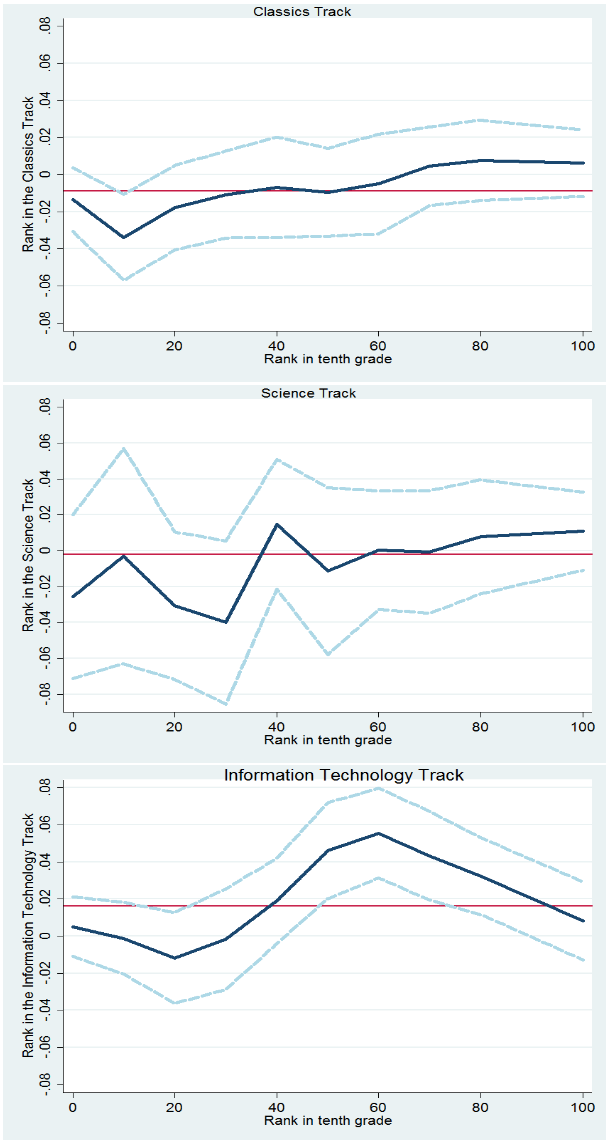
Note: The estimated effect of feedback on the national rank in the twelfth grade by quality of school at each decile of students' GPA performance in the tenth grade and the associated 95 % confidence interval. The effect of feedback on students' national rank when they realise they are in schools worse than the average quality school (on the left) and better than the average quality schools (on the right). The national rank is calculated based on the five core educational subjects (incentivised). The regressions are conditional on the students' characteristics: gender, age, a dummy that takes the value of one if the student is early enrolled in school and dummies for the track each student chooses in the twelfth grade. Standard errors are clustered at the school level.

Figure 2.15: Treatment effects on the standardised score by school quality conditional on prior performance



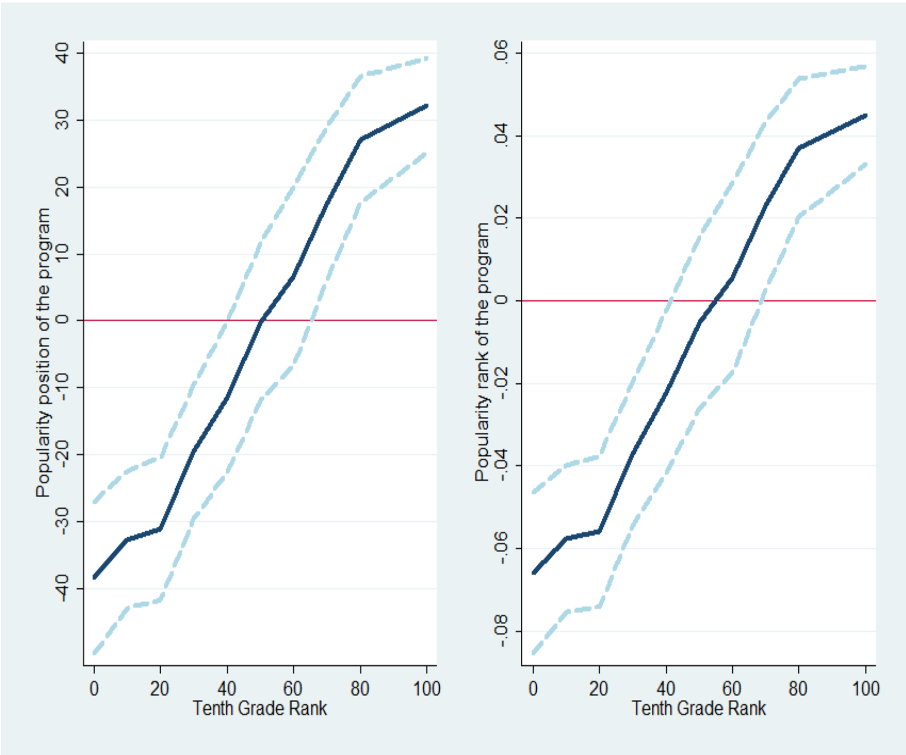
Note: The estimated effect of feedback on the standardised score in the twelfth grade by quality of school at each decile of students' GPA performance in the tenth grade and the associated 95 % confidence interval. The effect of feedback on students' standardised score when they realise they are in schools worse than the average quality school (on the left) and better than the average quality schools (on the right). The standardised score is calculated based on the five core educational subjects (incentivized). The standardised score has a mean of zero and a standard deviation of one in each year. The regressions are conditional on the students' characteristics: gender, age, a dummy that takes the value of one if the student is early enrolled in school and dummies for the track each student chooses in the twelfth grade. Standard errors are clustered at the school level.

Figure 2.16: Treatment effects on the rank nationwide by school quality conditional on prior performance



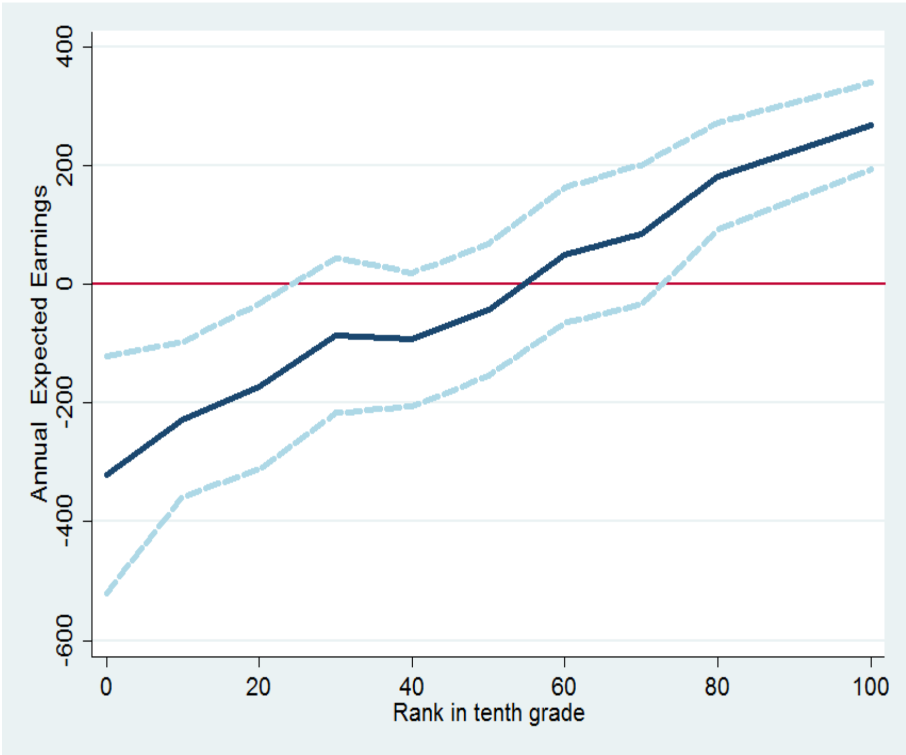
Note: The estimated effect of feedback on the national rank in the twelfth grade by track/specialisation at each decile of students' GPA performance in the tenth grade and the associated 95 % confidence interval. Three tracks are available in all schools: Classics, Science and Information Technology. The regressions are conditional on the students' characteristics: gender, age, a dummy that takes the value of one if the student is early enrolled in school and dummies for the track each student chooses in the twelfth grade. Standard errors are clustered at the school level.

Figure 2.17: Treatment effects on the popularity position and rank of the program admitted conditional on prior performance



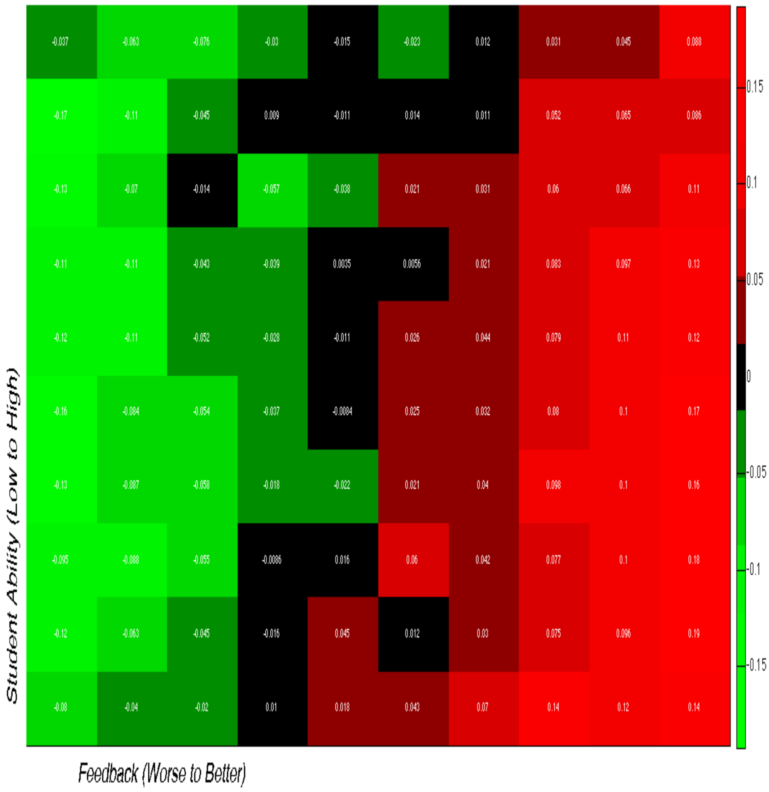
Note: The estimated effect of feedback on the popularity position (on the left) and rank(on the right) of the program admitted and the associated 95 % confidence interval. There are 672 programs in total. Popularity position and rank measured by the average University Department cut-off score over seven years. The regressions are conditional on the students' characteristics: gender, age, a dummy that takes the value of one if the student is early enrolled in school and dummies for the track each student chooses in the twelfth grade. Standard errors are clustered at the school level.

Figure 2.18: Treatment effects on the annual expected earnings conditional on prior performance



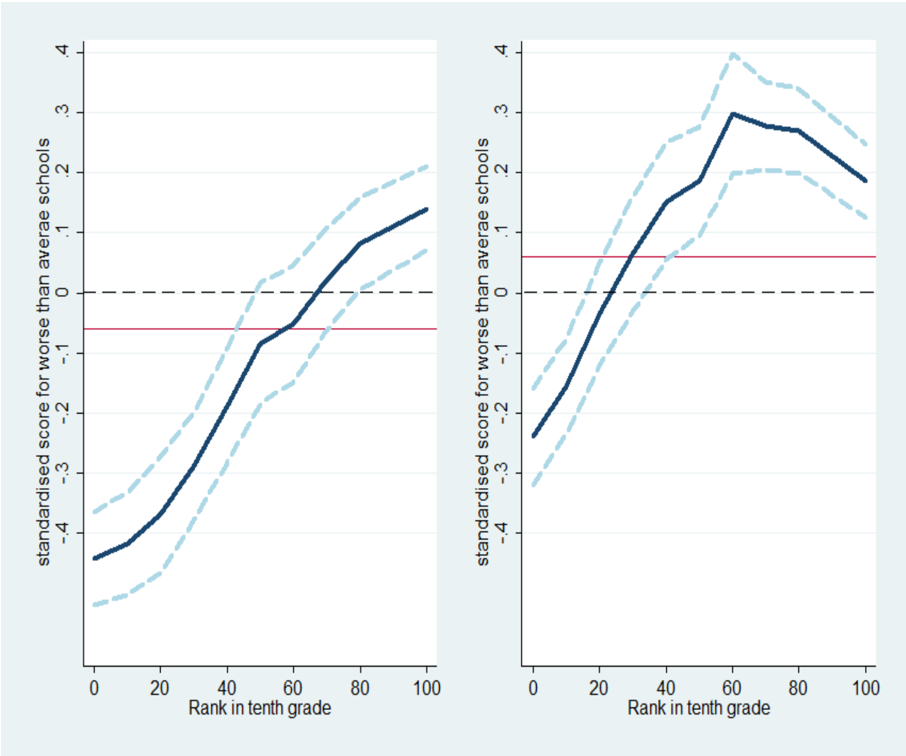
Note: The estimated effect of feedback on the expected annual wage at each decile of students' GPA performance in the tenth grade and the associated 95 % confidence interval. The annual expected earnings are calculated based on the actual annual earnings of older graduates who studied the same college field. The regressions are conditional on the students' characteristics: gender, age, a dummy that takes the value of one if the student is early enrolled in school and dummies for the track each student chooses in the twelfth grade. Standard errors are clustered at the school level.

Figure 2.19: Positive and Negative Surprise



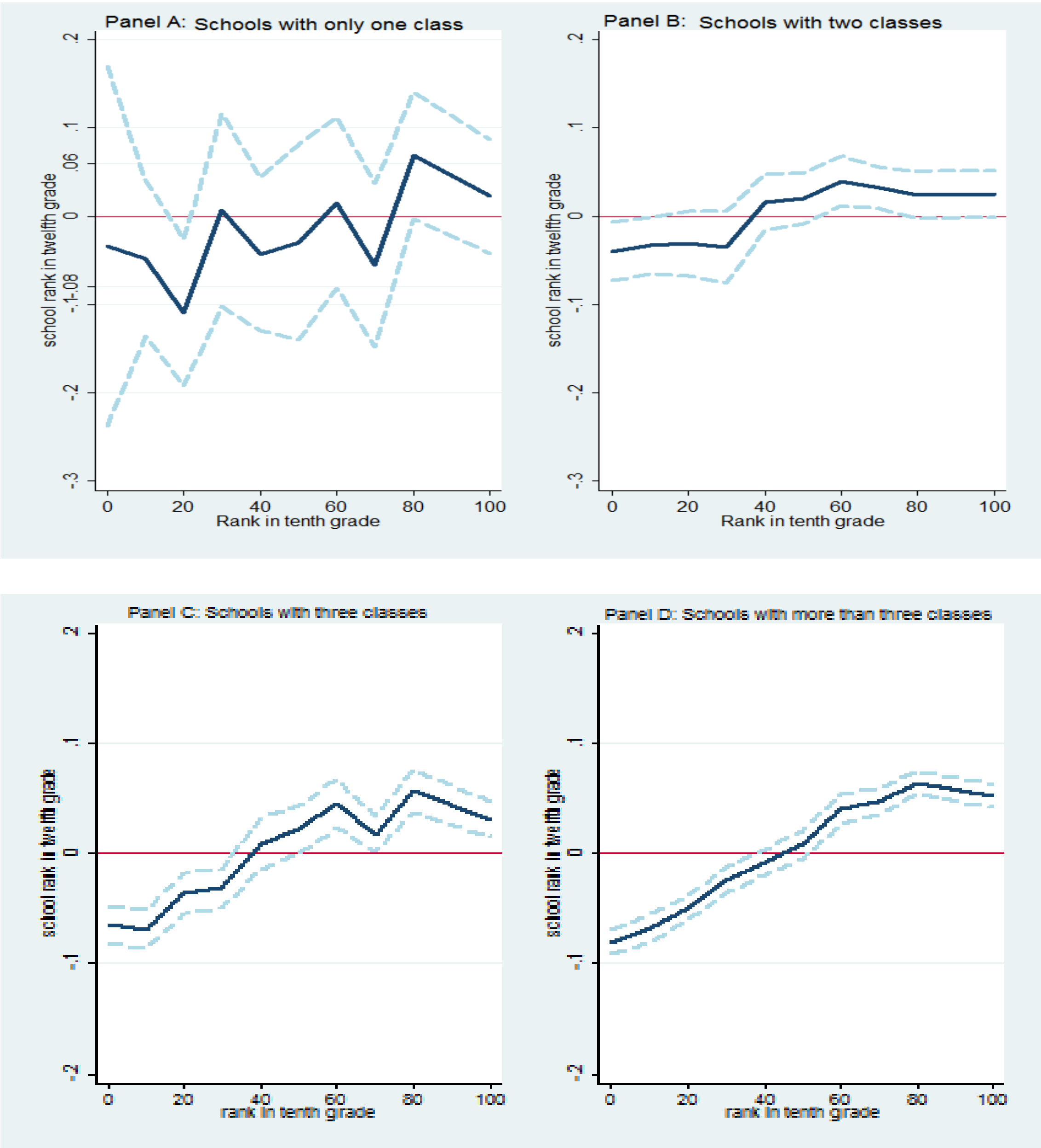
Note: Treatment effect for students with positive or negative surprise. Student performance (in deciles) in tenth grade on the vertical axis and student performance in eleventh grade (in deciles) on the horizontal axis.

Figure 2.20: Treatment effects on the standardised score by school quality conditional on prior performance



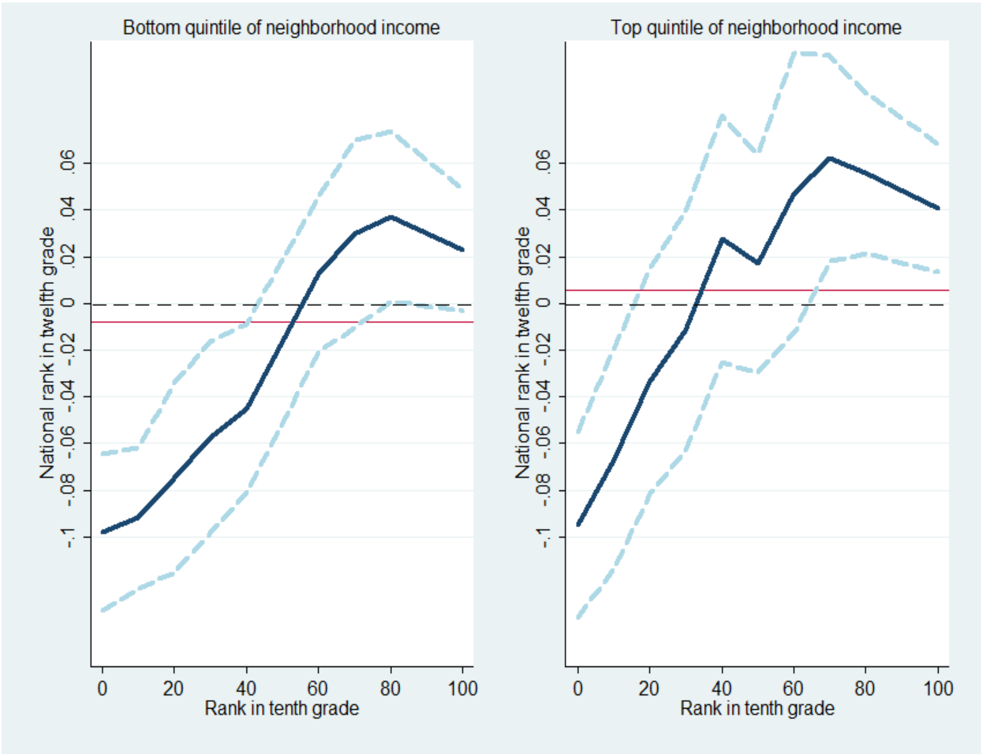
Note: The estimated effect of feedback on the standardised score in the twelfth grade by quality of school at each decile of students' GPA performance in the tenth grade and the associated 95 % confidence interval. The effect of feedback on students' standardised score when they discover they attend a school that is worse than the average quality school (on the left) and better than the average quality schools (on the right). The standardised score is calculated based on the five core-education subjects (incentivized). The standardised score has a mean of zero and a standard deviation of one in each year. The regressions are conditional on the students' characteristics: gender, age, a dummy that takes the value of one if the student is early enrolled in school and dummies for the track each student chooses in the twelfth grade. Standard errors are clustered at the school level.

Figure 2.21: Treatment effects on the rank within the school conditional on prior performance for different school cohorts' size



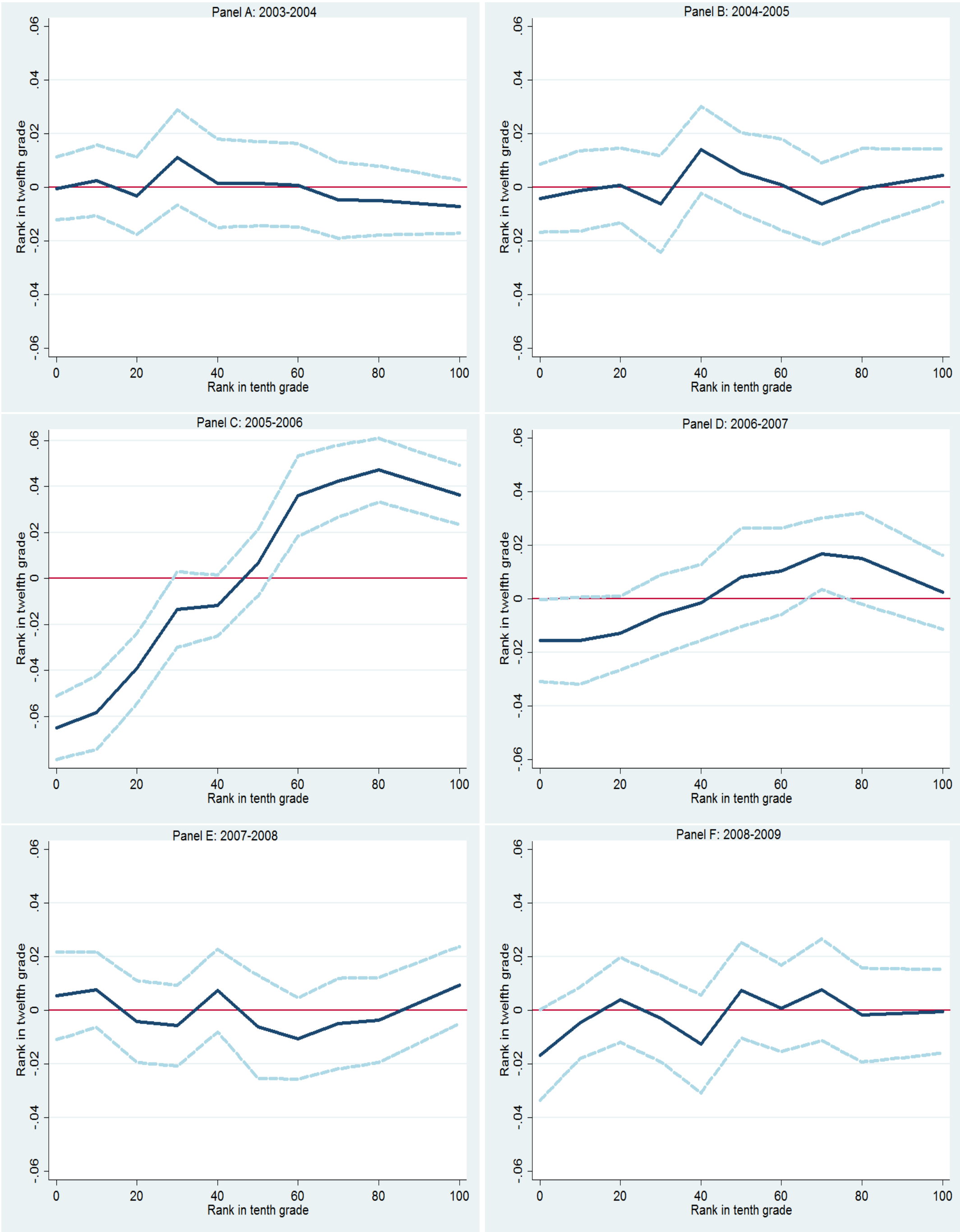
Note: The estimated effect of feedback on the school rank in the twelfth grade by capacity of school at each decile of students' GPA performance in the tenth grade and the associated 95 % confidence interval. The regressions are conditional on the students' characteristics: gender, age, a dummy that takes the value of one if the student is early enrolled in school and dummies for the track each student chooses in the twelfth grade. Standard errors are clustered at the school level.

Figure 2.22: Treatment effects on twelfth grade national rank for the bottom and top quintiles of neighborhood income



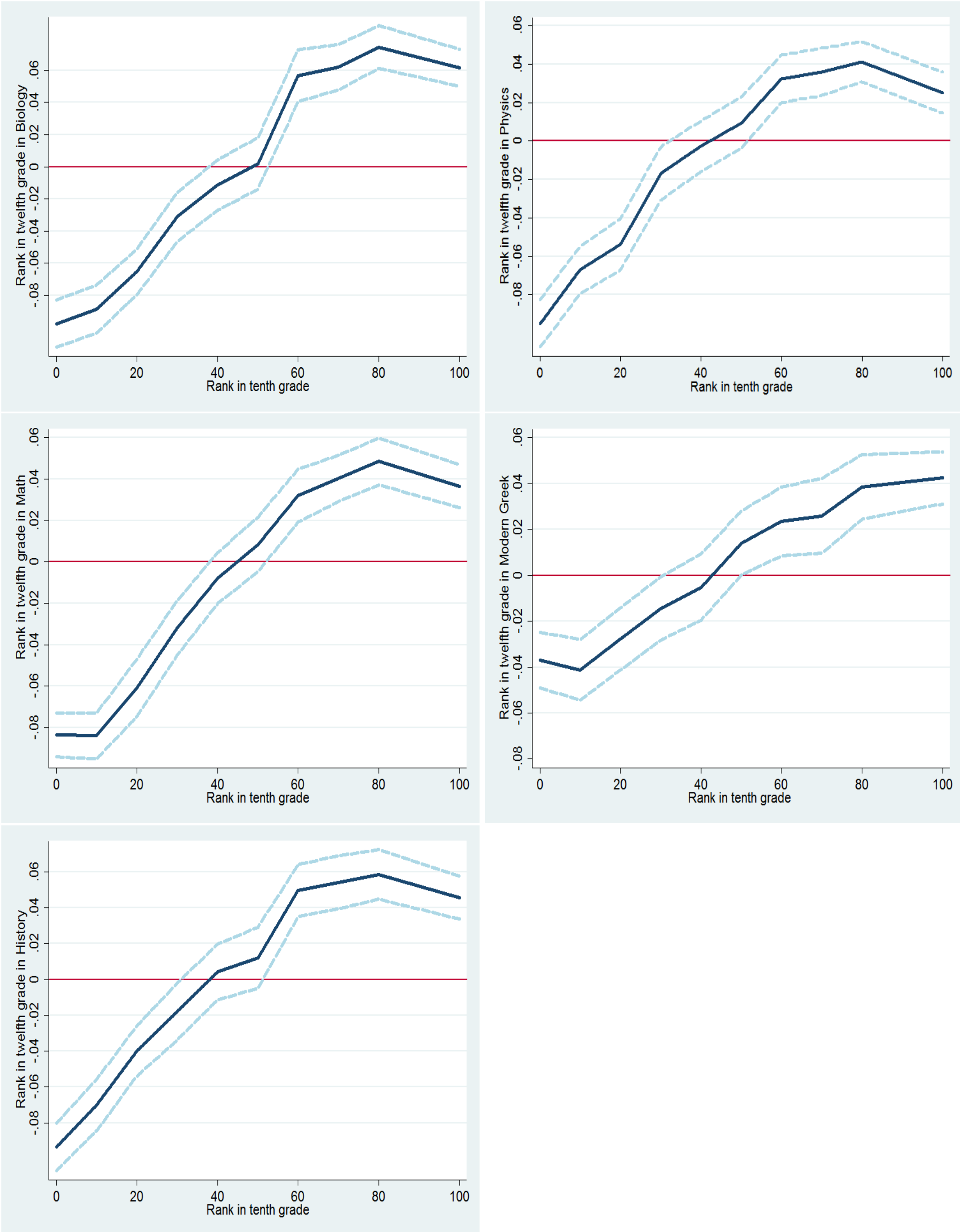
Note: The estimated effect of feedback on the national rank in the twelfth grade for the bottom (on the left) and top (on the right) quintiles of neighborhood income at each decile of students' GPA performance in the tenth grade and the associated 95 % confidence interval. The national rank is calculated based on the five core-education subjects (incentivized). The regressions are conditional on the students' characteristics: gender, age, a dummy that takes the value of one if the student is early enrolled in school and dummies for the track each student chooses in the twelfth grade. Standard errors are clustered at the school level.

Figure 2.23: Robustness checks



Note: Robustness checks: As if feedback was abolished in 2004 (Panel A), 2005 (Panel B), 2006 (Panel C), 2007 (Panel D), 2008 (Panel E) and 2009 (Panel F).

Figure 2.24: Feedback effects on twelfth grade rank nationwide for each subject separately conditional on prior performance



Note: The estimated effect of feedback on the twelfth garde rank nationwide at each decile of students' GPA performance in the tenth grade and the associated 95 % confidence interval. The regressions are conditional on the students' characteristics: gender, age, a dummy that takes the value of one if the student is early enrolled in school , school fixed effects and dummies for the track each student chooses in the twelfth grade. Standard errors are clustered at the school level.

2.9 List of Tables

Table 2.1: Descriptive Statistics

Variable	**Mean**	**Std. Dev.**	**Min.**	**Max.**
Student Characteristics				
Age	17.875	0.466	17	27
Early enrolment	0.167	0.373	0	1
Female	0.566	0.496	0	1
School cohort size	78.518	31.17	10	170
School GPA	85.930	10.186	49.44	100
National exam grade	62.843	19.362	7.550	98.857
Cohort size	63,186	8,710	50,061	71,796
logIncome(in 2009 Euro)	9.999	0.270	9.473	11.105
Retake the national exam	0.115	0.319	0	1
Specialty Characteristics				
Specialty:Classics	0.359	0.48	0	1
Specialty:Exact Sciences	0.164	0.371	0	1
Specialty:Information Technology	0.477	0.499	0	1
School Characteristics				
Private School	0.039	0.193	0	1
Experimental School	0.061	0.24	0	1
Public School	0.9	0.3	0	1
Urban	0.973	0.161	0	1
University Admission				
Admitted	0.823	0.381	0	1
College district different from school district	0.677	0.468	0	1
Number of university departments	8.293	10.543	1	242
Rank of admitted college in preference list	24.699	21.618	1	254
Places in tertiary education	60,960	6,268	52,450	68,136

Note: 45.842 obs. 7 cohorts. The variable "places in tertiary education" is calculated as the average across admitted students.

Table 2.2: Sample and Population

Variable	134schools Mean	1189schools Mean	Difference (b/s.e.)
Age	17.875	17.892	-0.017***
			(0.003)
Early enrolment	0.167	0.167	-0.0004
			(0.002)
Female	0.566	0.565	0.002
			(0.003)
School cohort size	78.518	75.358	3.160
			(0.197)
logIncome (in 2009Euro, annual)	9.999	9.938	0.060***
			(0.001)
Retake	0.115	0.112	0.003
			(0.002)
Specialty: Classics	0.359	0.366	-0.007
			(0.004)
Specialty: Exact Sciences	0.164	0.159	0.005
			(0.002)*
Specialty: Information Technology	0.477	0.475	0.002
			(0.003)
School and University Characteristics			
Private school	0.039	0.080	-0.041***
			(0.001)
Public schools	0.900	0.901	-0.001
			(0.002)
Experimental school	0.061	0.019	0.042***
			(0.001)
Urban	0.973	0.892	0.082***
			(0.002)
Admitted	0.823	0.803	0.020***
			(0.001)
Internal migration	0.677	0.800	-0.123***
			(0.002)
Rank of admitted college in preference list	8.293	8.584	-0.292***
			(0.065)
No of university departments in preference list	24.699	26.865	-2.166***
			(0.120)

Note: 45,842 obs. in sample and 431,469 obs. in population. There are in total 1,323 senior high schools in operation. Evening schools are excluded from the sample and the population

Table 2.3: Treatment and Control Group

Variable	Feedback Mean	No Feedback Mean	Difference (b/s.e.)
Student and Speciality Characteristics			
Age	17.835	17.909	0.074***
			(0.004)
Early enrolment	0.209	0.129	-0.080***
			(0.004)
Female	0.553	0.579	0.026***
			(0.005)
School cohort size	88.083	70.030	18.053***
			(0.288)
logIncome (in 2009Euro,annual)	9.988	10.005	0.017***
			(0.003)
Retake	0.104	0.124	0.020***
			(0.003)
Specialty: Classics	0.344	0.377	0.033***
			(0.004)
Specialty: Exact Sciences	0.176	0.154	-0.022***
			(0.004)
Specialty: Information Technology	0.480	0.469	-0.011**
			(0.005)
School and University Characteristics			
Private school	0.037	0.037	0.0003
			(0.002)
Public schools	0.905	0.897	-0.008
			(0.005)
Experimental school	0.058	0.066	0.007
			(0.005)
Urban	0.972	0.974	0.002
			(0.002)
Admitted	0.836	0.814	-0.022***
			(0.004)
Internal migration	0.475	0.635	-0.160***
			(0.002)
Rank of admitted college in preference list	9.657	7.068	-2.589***
			(0.115)
No of university departments in preference list	26.946	22.724	-2.589***
			(0.011)

Note: 21.965 obs. in treatment group and 23.781 obs. in control group. The feedback period is the pooled period from 2003 to 2005 while the non-feedback period consists of the pooled period from 2006 to 2009.

Table 2.4: Estimation results: Rank nationwide

Dependent Variable: Rank nationwide in incentivized subjects			
	Specifications		
Variable	(1)	(2)	(3)
Feedback*quintile5	0.042***	0.043***	0.045***
	(0.004)	(0.006)	(0.004)
Feedback*quintile4	0.036***	0.037***	0.040***
	(0.004)	(0.005)	(0.004)
Feedback*quintile2	-0.045***	-0.045***	-0.038***
	(0.004)	(0.004)	(0.005)
Feedback*quintile1	-0.088***	-0.088***	-0.079***
	(0.004)	(0.004)	(0.004)
Feedback	0.009***	0.009	-0.001
	(0.003)	(0.009)	(0.003)
quintile5	0.234***	0.235***	0.251***
	(0.003)	(0.004)	(0.004)
quintile4	0.094***	0.094***	0.102***
	(0.003)	(0.003)	(0.003)
quintile2	-0.081***	-0.083***	-0.093***
	(0.003)	(0.003)	(0.003)
quintile1	-0.176***	-0.177***	-0.192***
	(0.003)	(0.003)	(0.003)
Female	-0.008***	-0.008***	-0.011***
	(0.001)	(0.002)	(0.002)
Specialty: Science	0.055***	0.052***	0.048***
	(0.002)	(0.002)	(0.002)
Specialty: Classics	-0.022***	-0.021***	-0.021***
	(0.002)	(0.002)	(0.002)
Log Income	0.055***		
	(0.003)		
Experimental school	0.029***		
	(0.004)		
Private school	0.145***		
	(0.004)		
Urban	0.021***		
	(0.004)		
Year FE.	no	no	yes
School FE.	no	yes	yes
Observations	45,746	45,746	45,746
R squared	0.635	0.666	0.675
No of schools	134	134	134

Note: A constant is also included. Clusters at school level. *,**,*** denotes significance at the 10%,5% and 1% level respectively.

Table 2.5: Rank within the school in incentivized and non-incentivized subject

Dependent Variable: School Rank in incentivized and non-incentivized subjects				
	Incentivized subjects		**Non-Incentiv. subjects**	
Variable	(1)	(2)	(3)	(4)
Feedback*quintile5	0.045***	0.045***	0.005	0.005
	(0.004)	(0.004)	(0.006)	(0.006)
Feedback*quintile4	0.040***	0.040***	-0.005	-0.005
	(0.004)	(0.004)	(0.006)	(0.006)
Feedback*quintile2	-0.038***	-0.038***	-0.004	-0.003
	(0.004)	(0.005)	(0.006)	(0.006)
Feedback*quintile1	-0.079***	-0.079***	0.005	0.005
	(0.004)	(0.004)	(0.005)	(0.006)
Feedback	0.001	-0.001	0.003	0.001
	(0.003)	(0.003)	(0.004)	(0.004)
quintile5	0.251***	0.251***	0.256***	0.256***
	(0.004)	(0.004)	(0.005)	(0.004)
quintile4	0.102***	0.102***	0.103***	0.105***
	(0.003)	(0.003)	(0.003)	(0.004)
quintile2	-0.093***	-0.093***	-0.094***	-0.095***
	(0.003)	(0.003)	(0.004)	(0.005)
quintile1	-0.193***	-0.192***	-0.200***	-0.200***
	(0.003)	(0.003)	(0.004)	(0.006)
Female	-0.009***	-0.011***	0.054***	0.054***
	(0.001)	(0.002)	(0.003)	(0.002)
Specialty: Science	0.047***	0.048***	0.033***	0.034***
	(0.001)	(0.002)	(0.002)	(0.003)
Specialty: Classics	-0.019***	-0.021***	0.097***	0.097***
	(0.001)	(0.002)	(0.002)	(0.003)
Log Income	0.051***	0.049***	0.007	0.007
	(0.0004)	(0.001)	(0.004)	(0.004)
Experimental school	-0.041***	-0.038***	-0.003	-0.004
	(0.002)	(0.004)	(0.004)	(0.003)
Private school	-0.003	-0.004	0.030	0.032
	(0.003)	(0.004)	(0.016)	(0.018)
Urban	-0.017***	-0.016***	-0.003	-0.004
	(0.003)	(0.003)	(0.003)	(0.003)
Year FE.	no	yes	no	yes
Observations	45.746	45.746	45.746	45.746
R squared	0.674	0.675	0.542	0.543
No of schools	134	134	134	134

Note: Standard errors are clustered at the school level. A constant is also included. *,**,*** denotes significance at the 10%,5% and 1% level respectively.

Table 2.6: Estimation results : Differential Response by Gender

Dependent Variable: Rank in twelfth grade				
	Rank within the school		**Rank nationwide**	
Variable	(1)	(2)	(3)	(4)
Female*Feedback	-0.028	-0.028	-0.027	-0.027
	(0.005)***	(0.005)***	(0.005)***	(0.005)***
Female	0.054	0.054	0.052	0.052
	(0.003)***	(0.003)***	(0.003)***	(0.003)***
Feedback	0.009	0.002	0.009	0.008
	(0.003)***	(0.004)	(0.008)	(0.009)
Speciality in Science	0.198	0.199	0.198	0.196
	(0.004)***	(0.004)***	(0.004)***	(0.004)***
Speciality in Classics	-0.040	-0.039	-0.039	-0.040
	(0.003)***	(0.003)***	(0.005)***	(0.003)***
Income	-0.0001	-0.0001	0.0002	0.0002
	(0.0001)	(0.0001)	(0.0001)***	(0.0001)***
Private	-0.015	-0.015	0.134	0.134
	(0.008)*	(0.008)*	(0.016)***	(0.017)***
Experimental	-0.015	-0.015	0.017	0.017
	(0.006)**	(0.006)**	(0.018)	(0.018)
urban	-0.029	-0.029	0.007	0.007
	(0.007)***	(0.007)***	(0.015)	(0.015)
R^2	0.14	0.14	0.16	0.16
N	45,746	45,746	45,746	45,746
Year FE		✓		✓
No of schools	134	134	134	134

Note: Standard errors are clustered at the school level. A constant is also included. *,**,*** denotes significance at the 10%,5% and 1% level respectively. The rank in the twelfth grade here takes into account only the incentivized subjects. It is calculated within the school for columns (1) and (2) and across schools in columns (3) and (4)

Table 2.7: Capacity of schools

Variable	Mean	Std. Dev.	Min.	Max.
Schools with one class				
Public	0.899	0.302	0	1
Private	0.101	0.301	0	1
Experimental	0	0	0	0
Urban	0.378	0.485	0	1
Class size	18.130	5.717	10	29
No of schools	14			
No of students	522			
Schools with two classes				
Public	0.932	0.252	0	1
Private	0	0	0	0
Experimental	0.068	0.252	0	1
Urban	0.378	0.485	0	1
Class size	16.000	4.739	10	27
No of schools	38			
No of students	3,709			
Schools with three classes				
Public	0.941	0.235	0	1
Private	0.053	0.223	0	1
Experimental	0.006	0.077	0	1
Urban	0.986	0.115	0	1
Class size	18.211	4.998	10	32
No of schools	63			
No of students	9,959			
Schools with three classes				
Public	0.881	0.324	0	1
Private	0.035	0.184	0	1
Experimental	0.084	0.277	0	1
Urban	1	0	0	1
Class size	20.072	6.973	10	33
No of schools	74			
No of students	26,354			

Note: 111 senior high schools provided us with the eleventh and twelve grade classroom information. The number of classes in a school may not be stable across years. Some schools may expand and some others may shrink in some years.

Table 2.8: Loss of human capital in terms of labour force participants

Year	**Students Retaking**	**Potential Impact on Labour Market**
2003	7925	0.167%
2004	7223	0.150%
2005	6387	0.131%
2006	10421	0.213%
2007	6642	0.135%
2008	5730	0.116%
2009	4576	0.092%
2010	7680	0.153%

Table 2.9: Decision to Retake and Feedback

Dependent Variable: Repeat the national exams				
	LPM		**Probit**	**Logit**
Variable	(1)	(2)	(3)	(4)
Feedback* Misplacement	0.058	0.059	0.345	0.602
	(0.016)***	(0.016)***	(0.092)***	(0.181)***
Feedback	0.012	0.019	0.070	0.131
	(0.006)*	(0.007)**	(0.036)*	(0.074)*
Misplacement	-0.014	-0.015	-0.071	-0.099
	(0.014)	(0.015)	(0.077)	(0.142)
Age	-0.014	-0.019	-0.076	-0.157
	(0.003)***	(0.006)***	(0.039)*	(0.062)**
Early Enrolled	-0.005	-0.006	-0.011	-0.033
	(0.008)	(0.008)	(0.022)	(0.082)
Female	-0.007	-0.007	-0.044	-0.073
	(0.003)*	(0.004)*	(0.020)*	(0.038)*
Specialization in Classics	-0.020	-0.018	-0.113	-0.200
	(0.004)***	(0.007)*	(0.024)***	(0.046)***
Specialization in Science	0.013	0.016	0.090	0.169
	(0.005)**	(0.004)***	(0.026)***	(0.049)***
District Unemployment	0.005	0.002	0.025	0.046
	(0.002)**	(0.002)	(0.012)*	(0.019)**
If admitted in first place	-0.212	-0.218	-1.041	-1.964
	(0.008)***	(0.008)***	(0.035)***	(0.070)***
Internal Migration	0.064	0.072	0.445	0.889
	(0.005)***	(0.005)***	(0.037)***	(0.077)***
logIncome	-0.009			
	(0.011)			
Urban	0.024			
	(0.013)*			
Private	-0.056			
	(0.007)**			
Public	-0.039			
	(0.009)***			
R^2 or pseudo-R squared	0.05	0.06	0.07	0.07
Log likelihood			-13,432	-13,439
School FE		✓	✓	✓
Year FE	✓	✓	✓	✓
N	45,746	45,746	45,746	45,746

Note: A constant is also included. Standard errors are clustered at the school level. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Table 2.10: Decision to Retake, Feedback and Misplacement

Dependent Variable: Repeat the national exams				
	LPM		Probit	Logit
Variable	(1)	(2)	(3)	(4)
Feedback	-0.031 (0.007)***	-0.002 (0.008)	-0.007 (0.047)	-0.017 (0.090)
Feedback* Misplacement Quintile 5	0.045 (0.010)***	0.040 (0.009)***	0.219 (0.050)***	0.412 (0.095)***
Feedback* Misplacement Quintile 4	0.023 (0.010)**	0.023 (0.009)**	0.120 (0.049)**	0.231 (0.095)**
Feedback* Misplacement Quintile 2	0.004 (0.010)	0.007 (0.011)	0.049 (0.054)	0.103 (0.101)
Feedback* Misplacement Quintile 1	-0.034 (0.010)***	-0.031 (0.010)***	-0.151 (0.052)***	-0.274 (0.098)***
Misplacement Quintile 5	-0.017 (0.007)**	-0.018 (0.007)**	-0.103 (0.038)***	-0.184 (0.073)**
Misplacement Quintile 4	-0.025 (0.007)***	-0.025 (0.007)***	-0.139 (0.038)***	-0.262 (0.072)***
Misplacement Quintile 2	0.017 (0.007)**	0.016 (0.008)**	0.076 (0.039)*	0.143 (0.073)**
Misplacement Quintile 1	0.030 (0.007)***	0.031 (0.009)***	0.148 (0.043)***	0.273 (0.080)***
Female	-0.010 (0.004)***	-0.010 (0.004)***	-0.056 (0.020)***	-0.105 (0.037)***
Age	0.002 (0.007)	-0.001 (0.007)	-0.001 (0.035)	-0.002 (0.067)
Early Enrolled	0.011 (0.008)	0.009 (0.008)	0.047 (0.041)	0.087 (0.078)
Unemployment	0.005 (0.001)***	0.002 (0.002)	0.010 (0.011)	0.020 (0.021)
Internal migration	-0.024 (0.007)***	-0.022 (0.008)***	-0.109 (0.038)***	-0.211 (0.075)***
Specialization in Science	-0.007 (0.005)	-0.004 (0.005)	-0.018 (0.025)	-0.036 (0.048)
Specialization in Classics	-0.018 (0.004)***	-0.017 (0.004)***	-0.093 (0.024)***	-0.175 (0.045)***
Private	-0.087 (0.011)***			
Public	-0.040 (0.009)***			
LogIncome	-0.033 (0.008)***			
Urban	0.006 (0.010)			
R^2 or pseudo-R squared	0.03	0.04	0.06	0.06
Log likelihood			-14,062	-14,063
School FE		✓	✓	✓
Year FE	✓	✓	✓	✓
N	45,746	45,746	45,746	45,746

Note: A constant is also included. Standard errors are clustered at the school level. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Table 2.11: Drop out rate and Transfers

	Drop out 10 11	**Transfers 10 11**	**Drop out 11 12**	**Transfers 11 12**
	%	%	%	%
2000-2001	8.89			
2001-2002	12.31	8.89		
2002-2003	11.07	8.4	6.07	7.02
2003-2004	8.87	7.93	5.14	9.47
2004-2005	10.71	7.46	6.67	6.45
2005-2006	9.41	6.35	5.67	6.58
2006-2007	12.59	9.46	6.15	8.60
2007-2008	11.71	5.85	8.63	6.80
2008-2009	13.71	7.52	6.13	6.58
2009-2010	10.56	6.45	6.01	8.38
2010-2011	9.92	5.76	6.19	8.61

Note: The first column shown the percentage of students who drop out from school between the tenth and eleventh grade. The second column shown the percentage of students who transfer to a school in the eleventh grade. The third column shown the percentage of students who drop out from school between the eleventh and twelfth grade. The fourth column shown the percentage of students who transfer to a school in the twelfth grade. Data from 134 schools are used.

Table 2.12: Estimation results : Drop out

Dependent Variable: Dummy for drop out		
	Specifications	
Variable	(1)	(2)
Feedback*quintile5	0.009	0.010
	(0.007)	(0.007)
Feedback*quintile4	0.007	0.007
	(0.007)	(0.007)
Feedback*quintile2	0.009	0.010
	(0.008)	(0.008)
Feedback*quintile1	0.013	0.014
	(0.015)	(0.016)
Feedback	0.017	0.041
	(0.019)	(0.033)
quintile5	0.000	-0.001
	(0.004)	(0.004)
quintile4	-0.006	-0.006
	(0.005)	(0.005)
quintile2	0.025***	0.025***
	(0.006)	(0.006)
quintile1	0.153***	0.153***
	(0.003)	(0.014)
Female	-0.011***	-0.011***
	(0.003)	(0.004)
Absences10	0.001***	0.002***
	(0.0001)	(0.0001)
Year FE.	no	yes
Observations	56.041	56.041
R squared	0.130	0.203
No of schools	134	134

Note: A constant is also included. Clusters at school level. *,**,*** denotes significance at the 10%,5% and 1% level respectively. Quintiles are constructed based on the school performance in tenth grade used.

Table 2.13: Estimation results: Different outcome variables

Dependent Variable: Rank in twelfth grade			
	Specifications		
Variable	(1)	(2)	(3)
Feedback*quintile5	0.026***	0.030***	0.050***
	(0.004)	(0.007)	(0.005)
Feedback*quintile4	0.022***	0.015***	0.032***
	(0.004)	(0.007)	(0.005)
Feedback*quintile2	-0.029***	-0.032***	-0.042***
	(0.004)	(0.007)	(0.005)
Feedback*quintile1	-0.052***	-0.045***	-0.066***
	(0.004)	(0.006)	(0.005)
Feedback	0.002	0.008	-0.0004
	(0.003)	(0.005)	(0.003)
quintile5	0.257***	0.247***	0.245***
	(0.003)	(0.005)	(0.003)
quintile4	0.109***	0.110***	0.107***
	(0.003)	(0.005)	(0.003)
quintile2	-0.097***	-0.100***	-0.091***
	(0.003)	(0.005)	(0.003)
quintile1	-0.207***	-0.231***	-0.210***
	(0.003)	(0.005)	(0.003)
Female	-0.019***	0.030***	-0.014 ***
	(0.001)	(0.002)	(0.001)
Early Enrollment	0.010***	0.009***	0.011 ***
	(0.002)	(0.002)	(0.002)
Specialty: Science	0.006***	0.019*** *	0.023***
	(0.002)	(0.004)	(0.002)
Specialty: Classics	0.010***	0.098***	-0.059
	(0.002)	(0.003)	(0.002)
Observations	45,746	45,746	45,746
R squared	0.661	0.674	0.625
No of schools	134	134	134

Note: A constant is also included. The outcome in the first column is the rank calculated based on the five core subjects and the four Track subjects. The outcome in the second column is the rank in Modern Greek. The outcome variable in the third column is calculated based on five subjects in the feedback regime and two subjects in the non-feedback regime. Standard errors clustered at the school level. Year fixed effects included. Clusters at school level. *,**,*** denotes significance at the 10%,5% and 1% level respectively.

Table 2.14: Descriptive Evidence of Social Mobility

	Quintiles of Neighborhood Income									
	Quintile1		**Quintile2**		**Quintile3**		**Quintile4**		**Quintile5**	
Quintiles of program' popularity	Feed.	No Feed.	Feed.	No Feed.	Feed.	No Feed.	Feed.	No Feed.	Feed.	No Feed.
Quintile1	4.6	3.4	4.2	3.2	4.5	2.9	2.7	2.1	3.1	2.7
	+1.2		+1		+1.6		+0.6		+0.4	
Quintile2	3.9	3.5	3.5	3.1	4.1	3.5	2.5	2.7	3.1	3.2
	+0.4		+0.5		+0.6		-0.2		-0.3	
Quintile3	3.6	3.4	3.2	3.2	3.8	3.5	2.7	3	3.2	3.4
	+0.2		0		+0.3		-0.3		-0.3	
Quintile4	3.1	3.0	2.8	3.2	3.5	4.3	2.5	3.2	3.4	3.8
	+0.1		-0.4		-0.8		-0.7		-0.4	
Quintile5	2.9	2.6	2.6	3.0	3.9	4.1	2.8	3.2	3.6	4.2
	+0.3		-0.42		-0.2		-0.4		-0.6	
Total	18.1	15.9	16.3	15.7	19.8	18.3	13.2	14.2	16.3	17.4
	2.2		**+0.6**		**+1.5**		**-1**		**-1.1**	

Note: The variable "Quintile 1" represents the bottom quintile of program's popularity and neighbourhood income. The "Quintile 5" denotes the top quintile of the program's popularity and neighbourhood income. For each quintile of neighbourhood income two percentages are reported: the first one corresponds to the feedback period and the second one to the non-feedback period. The differences between the percentage in the feedback period and the non-feedback period for each quintile of program's popularity are also reported.

Chapter 3

Social Interactions Through Space and Time: Evidence from college enrolment

3.1 Introduction

In the recent years the literature on the role of social interactions in economic behavior has expanded rapidly. This doesn't come as surprise when one thinks the importance of those effects in every day decision-making. The basis of decision-making though in almost every context is information. Humans are social beings and we naturally collect information through social interactions in order to inform our goals and choices. This is even more pronounced among adolescents. In developmental science, it has been widely argued that adolescents and young adults regularly mimic the choices and behavior of role models in their environment (Bell 1970).

Brock and Durlauf 2001 define social interactions as the idea that an individual's marginal utility with respect to other individuals' choices in his reference group is positive. The desire to conform induces prevalent patterns of behavior even among agents with heterogeneous tastes over externalities from other individuals' choices (Bernheim 1994). Social interactions within a reference group have been shown to affect students' achievement. However, there is little evidence on the effect of social interactions on the decisions of college enrollment and academic mobility. Moreover, social interactions can explain variation in choices across groups with similar characteristics. For example, Schelling 1973 provide early evidence of social interactions in binary choice in a profusion of contexts such as driving style and athletic play. Intuitively, conformity causes social interactions to be interconnected with neighborhood effects. Physical proximity amplifies the interplay of utility spillovers from other agents' choices and the combined effect becomes area specific. In an educational context, Garner and Raudenbush 1991 provide evidence of a positive relation between neighborhood quality and educational attainment.

There is evidence that peers' decision affect scholastic performance in elementary, middle and high school but also during college. Hoxby 2000a examines the effect of social interaction in grade school and finds that students who were randomly assigned to classes with students who have high reading scores relative to the school and grade, received higher reading scores. Hanushek et al. 2003 find that peer achievement has a positive effect on achievement growth. In particular, 0.1 standard deviation increase in peer average achievement leads to a 0.02 increase in student's performance. Zimmerman 2003b examines the effect of social interaction using freshmens' SAT score. He finds strong positive social interaction effects among roommates at almost all parts of the ability distribution. Cipollone and Alfonso 2007 find strong social interactions inter alia the decision to stay longer in school. When men were exempted from the compulsory military services -due to an earthquake- and stayed longer in school, the graduation rates of young women in the affected areas rose by about 2 percentage points. Fletcher 2006 using survey data, finds strong evidence of social interactions college preferences and college enrollment. Giorgi et al. 2007 find that ones' behavior influences the educational decision while in college, indicating the importance of social interaction even at a later stage of someone's academic life. Sacerdote 2011 examines social interaction effects at the room and accommodation level where students are randomly assigned. He does not find any significant influence of peers.

In this paper we examine the effect of social interactions on the decisions of adolescents and young adults regarding college enrollment and academic mobility. We use a new dataset from Greece that contains information on exam scores, college enrollment and educational mobility for every student in six cohorts. We exploit the particular institutional setting in Greece, in which schools are build very close to each other. This setting allows for rich variation of school characteristics within a relatively contained geographical area. We exploit this exogenous variation in group characteristics over time and space to address the endogenous nature of the social interaction groups. The social interaction effects are defined as contextual interactions that induce different mappings from individual characteristics to outcomes (Bryk and Raudenbush 2001). Reference groups are viewed as ecologies in which the social backgrounds affect individual choices of otherwise similar agents (Raudenbush and Sampson 1999).

Similar age peers in one's vicinity consist a natural reference group that provide valuable and otherwise costly information, necessary in academic decision making. We widen the reference group and examine social interactions with respect to a series of reference groups: same-cohort school peers, different-cohort school peers, same-cohort peers in the neighborhood and different-cohort peers in the neighborhood.

There are particular advantages in having the universe of high school graduates for a country. First, we can observe the behaviour of all students regarding their education decisions and not only of some groups of students. Second, we are able to observe different reference groups. A student may be affected by the decisions of same age or older peers in his school and neighborhood. We contribute to the literature by comparing the size of the social interaction effects across distance in space and age.

Empirical analysis of social interactions on students' decisions has been open to question because of the difficulties in disentangling these effects from other confounding influences.[1]. We use an instrumental variable approach and we exploit spatial and cohort-to-cohort variation to combat potential endogeneity problems and the well known "reflection-problem" (Manski 1993, Manski 2000). There are two sources of potential endogeneity: Self selection into social groups and common shocks that affect every member of a social group. Reflection may arise from reverse causality because the outcomes of members in the same groups and their decisions are simultaneous. In other words, it is difficult to disentangle if one's actions are the cause or the effect of his peers' influence. These challenges are standard in the social interactions literature. The institutional setting behind our study refrains students from endogenously select their peers in school, facilitating the validity of the identification strategy. Moreover, the geographical density of schools allows us to define social groups wider than a student's schoolmates. Motivating from the idea of role modelship, we battle the simultaneity challenge by investigating social interactions between peers in consecutive cohorts.

By using multiple cohorts and conditioning on school and neighbourhood fixed effects as well as school-and neighborhood- specific time trends we are able to control for unobserved time-varying factors that might confound peer effects in schools and neighbourhoods. We use an instrumental variable approach to combat endogeneity and reflection. We show that within schools and neighbourhoods, there is considerable cohort-to-cohort variation in the proportion of female students that can be attributed to random factors.

We find positive spillover effects between one's decision to enrol in college and that of their peers. More specifically, the results found here indicate that students who attend a high school with a hundred percent more schoolmates who enrol in college are 12.6 percent more likely to attend college. We also find positive spillovers regarding the decision of educational mobility. Students are 10.7 percent more likely to move to a different city to study if their older peers in school do so, a hundred percent more often. We find that these externalities decrease with the size of reference group.

The policy implications of social interactions can be indirect. The skills and resources that characterise a reference group are usually fixed. As a consequence, an improvement in someone's group characteristics means an equivalent deterioration in someone else's group attributes. Some may argue that the redistribution in favor of disadvantaged students can act as a boost in their scholastic outcomes, when the redistribution comes from more advantaged areas where students might depend less

[1] The existing literature that deals with identification of the social comparison effects use either laboratory experiments (Armin and Andrea 2006), natural experiments (Zimmerman 2003b), quasi-experimental designs (Hoxby 2000a), or fixed effects (Hanushek et al. 2003

on their peers' quality. For example, Arcidiacono and Nicholson [2005] suggest that the existence of social interaction effects supports claims against school vouchers. This is because, the best students leaving public schools can be detrimental to the students left behind.

The paper is organized as follows. Section 3.2 describes the unique dataset used and the institutional setting related to college admission. The empirical strategy used to identify social interactions is analysed in Section 3.3. Section 3.4 discuss the validity of the identification strategy. We present and discuss the results in college enrollment and educational mobility in Section 3.5. Finally, Section 3.6 concludes.

3.2 Data and Institutional Setting

3.2.1 How are students admitted to college

The transition from high school to post-secondary education in Greece is based on an unusually systematic and transparent allocation of students to university departments[2] In particular, every high school student who completes the twelfth grade receives an admission score, which is the only criterion for university admission and weights: (i) her performance in national twelfth grade exams [3] (ii) her grade twelve within school performance which is a combined score for homework and midterm exams in each subject.

After receiving their admission scores, students are required to submit a list of ranked choices of specific departments in universities that are relevant to their twelve grade track. For example, students outside the Classics track cannot list Law schools. Each university department generally offers one major of bachelor degree and no minor specializations can be declared. Every university department admits a pre-specified number of students. A computerized system at the Ministry of Education ranks students by their admission score and assigns the highest ranked student to her preferred choice. It then moves to the next student and assigns her

[2]Every tertiary education institute in Greece is public as free education is a constitutional right. Degrees awarded by private colleges are not recognized by the state.

[3]The twelfth grade exams are written exams administered nationally only once every year and last from late May to early June. The exams are proctored and marked externally. Exam markers do not observe the name, school, or even the city of the student whose paper they grade. Students usually take six component exams, with a combination of common subjects(Language, Mathematics, Physics, Biology or History) and four compulsory track-specific subjects and one elective exam. There are three tracks: Classics, Natural Sciences and Technical Studies. The overall score is the unweighted average of these scores. Students who fail are allowed to retake the exam the next year. In addition, students are not allowed to take the national exams early.

to the first department in her list in which there is an available place, and so and so forth. In this context, students have incentives to truthfully reveal their preferences.

University departments must enrol the students assigned to them by the Ministry of Education. The Ministry of Education announces the score of the last admitted student in each university department. The last admitted students in more prestigious departments have generally higher scores in comparison to those in less prestigious ones. Once a student admitted they cannot transfer to a different major. College education is completely publicly funded and every student is exempted for college fees. Private donations to colleges are against the law.

3.2.2 Data

For the empirical analysis we construct a unique dataset of all students graduating from high school in Greece from 2003 to 2009. We obtain the information from various sources:

1. Administrative data from the Hellenic Ministry of Education containing course taking information and exam grades in the final year, gender, year of birth, graduation year and college admission information. In addition, the total number of places in tertiary education in each year is provided.

2. School specific information such as name of school, type of school (private, public[4], experimental[5]), geographical coordinates and name of preferacture it belongs to[6]. There are 1319 high schools in Greece[7].

3. The Ministry of Finance provided us with average net income information at the postcode of the school in 2009 Euro.

4. The Ministry of Internal Affairs provided us with urban density information. Urban areas are those with more than 20,000 inhabitants.

5. Geographical coordinates for every tertiary education institute in Greece. There are fifty five college campuses. Not all campuses offer the same majors.

[4]Students are assigned to public schools according to a school district system

[5]Admission to experimental schools is based on a lottery

[6]There are fifty two preferactures in Greece. Preferactures are classified by the Hellenic National Statistical Authority

[7]Of which, 112 are private, and 1207 public. Of those 1207 public schools, 23 are experimental. There are no private experimental schools in Greece. 74 evening high schools for employed people of usually older age are excluded from our analysis

The median distance of a school from each nearest neighbouring school is 0.32 miles.[8] We use cluster analysis to define and construct neighborhoods within a mile from each school. We construct 406 clusters that cover the whole country. Every cluster is a neighborhood that contains all twelve-grade students who attend any other high school within a mile (1.06 miles) radius from one's high school[9]. Figures 1 maps all high schools and tertiary education institutes in our dataset.

Our analysis uses information regarding characteristics and choices of older school peers. Because of this, we use data on student cohorts from 2004 to 2009[10] Furthermore, our discussion of academic mobility refers to the decisions of students to move to a different preferacture in order to study, given they were admitted to some college. Thus, for this part, we focus only on admitted students[11]. Lastly, we drop 35,808 obs. for which the group of schoolmates overlapped perfectly with the social group of their neighborhood. This exclusion allows us to compare spillover effects from social groups of various sizes. We consolidate our sample by dropping observations with missing values.

Table 3.1 describes our pooled data across cohorts. Fifty seven percent are females. Ninety percent of the students reside in urban areas. More than 90 % of schools are public. Although, mean postcode income among private schools is significantly higher compared to public schools, mean national exam score doesn't seem to differ much. Experimental schools are in more affluent areas in comparison to other public schools as revealed by their higher mean postcode income. The mean national exam score of students attending experimental schools is much higher than the score achieved by students in private or public schools. Each neighbourhood contains on average 4 schools and 929 student observations.

3.3 Empirical Strategy

We start off by defining one's reference group as his same-cohort school peers. We investigate the hypothesis that social or collective behaviour patterns drive individual preferences because agents derive utility from conformity or provide access to

[8]Mean of distance from nearest neighbour: 1.85 miles. Standard deviation: 18.37 miles. 25th percentile:0.07 miles. 75th percentile: 0.77 miles.

[9]We exploit the fact that many schools were built very close to each other in most urban settings in Greece. This is more prevalent in Attica, the region surrounding the city of Athens, the capital of Greece. To give an example, in the cartier of Grava in Athens, there are six high schools next to each other along with several elementary and middle schools that form a humongous school building complex. According to the 2001 census, Attica holds around 36 percent of the total population.

[10]The first cohort in our sample, 2003 (size: 59,102 obs.), is used as a reference group for the 2004 cohort.

[11]In the academic mobility analysis we exclude 60,356 students who did not enrol in college

information.

In particular, we investigate whether a student's decision to enrol in college depends on the decision of his peers in school by using the following regression:

$$IfEnrolled_{i,s,t} = \alpha + \gamma\overline{IfEnrolled_{i-1,s,t}} + \beta X_{i,s,t} + \kappa T_t + \mu S_s + \pi_s year + \epsilon_{ist} (1)$$

where $IfEnrolled_{i,s,t}$ takes the value one if student i in school s and year t enrols into college and $\overline{IfEnrolled_{i-1,s,t}}$ is the fraction of all other students except of student i in school s and year t, who enrol into college. So, we regress a student's i decision to enrol in college on the mean enrollment of his peers in school s in year t and other controls. Our covariates include a dummy for being female, a student's admission score, dummies for chosen track in the senior year of high school, dummies for the school each student attended, school specific time trends and year dummies. To control for time-varying unobserved factors that may be correlated with mean college enrolment we include a full set of school-specific linear time trends.

The main coefficient of interest is γ, which captures how the mean enrollment of one's school peers affects his decision to enrol in college. Initially, we employ ordinary least squares to estimate peer effects in education decisions. There are at least two sources of potential bias here: (1) endogeneity and (2) the reflection problem (Manski [1993], Manski [2000]).

Firstly, in many settings individuals self-select themselves into a specific group of peers that generates endogeneity issues if the variables that are responsible for this choice are not fully observable. Students who choose to attend the same high schools might share the same observed and unobserved characteristics. In this case, if we find a relationship between the observed characteristics and the outcome variable it might not be causal. This could be coming from the fact that unobserved characteristics might also affect the outcome variables. This potential unobserved heterogeneity that drives selection into social groups may bias our estimates. Nevertheless, self selection of students into schools is restricted in our setting because students are assigned to public schools[12] based on geographical criteria and they cannot choose their school peers endogenously, by construction. Therefore, social group membership is as good as random, since it does not depend on observables.

Endogeneity may also result from unobserved common group effects, such as teacher and school quality, that affect every student in a social group and render the identification of social interactions challenging. We contribute to the literature

[12] 92 % of students in our sample attend public or public experimental schools

by mitigating the endogeneity challenge that stems from common group shocks. We take advantage of a special institutional setting with rich spatial and over time variation in school characteristics. We use cluster analysis to construct geographical units wider than the school district; namely neighborhoods. Those geographical units are big enough to allow for school diversity but also compact enough to capture common behavioral patterns in the area. In particular, we exploit our special institutional setting to identify same-cohort peers who do not attend the same school. We identify same-cohort peers who live in 1 mile radius and attend different schools. We call this group of same-cohort peers who live very close to each other "neighbours". In addition to their same-cohort schoolmates, students are likely to interact with their same-cohort neighbours and they might also be affected by their decisions. A students' neighbours attend different schools and face different school environments. In each neighbourhood, there are students who attend on average four (4.449) different schools (Table 3.1). The basic idea here is to compare students' decisions from consecutive cohorts who have similar characteristics and face the same neighbourhood environment but attend different schools, but one cohort has more female students than the other. Thus, it becomes feasible to isolate the impact of a peer group from the impact of each student's school itself.

Second, reflection may arise because we cannot distinguish whether someone's action is the cause or the effect of his peers' outcomes. In other words, one's decision is simultaneous with that of his peers. We battle the simultaneity challenge by using as an IV the *time lagged* gender composition in the school and neighborhood level. So we compare the decisions of students from consecutive cohorts who had a different fraction of one-cohort-older peers in their school or neighbourhood who enrolled in college. To build some intuition here, peers one-cohort-older might provide information to younger peers about the costs and the benefits of attending college or migrating to another city or they might function as "role models".

Estimating equation (1) using OLS will lead to biased results. In order to address these concerns we propose the proportion of girls in someone's reference group in the previous period as a source of variation for mean enrollment in college. The intuition is that an individual's academic decision may be related with their gender, but not the gender composition of their environment. This satisfies the exclusion restriction for the validity of our instrument.

We control for unobserved characteristics of schools, students and neighbours that are correlated with the percentage of females and could also be correlated with students' performance. We do this by exploiting variation in the gender composition across consecutive cohorts within the same school and neighbourhood. By using multiple cohorts and controlling for school and neighbourhood fixed effects, we take into account unobserved factors that might invalidate the school and neighbourhood peer effects analysis.

The first stage regressions are:

$$\overline{IfEnrolled_{g,t}} = \phi_1 + \kappa_1 \overline{IfFemale}_{g,\mathbf{t}} + +\beta_1 X_{i,g,t} + T_t + S_g + \pi_{1g} year + e_{1,g,t} (2)$$

$$\overline{IfEnrolled_{g,t-1}} = \phi_2 + \kappa_2 \overline{IfFemale}_{g,\mathbf{t-1}} + +\beta_2 X_{i,g,t} + T_t + S_g + \pi_{2g} year + e_{2,g,t-1} (3)$$

$$g \in \{\{school\}, \{neighborhood\}\}$$

where $\overline{IfFemale}_{j,t}$ and $\overline{IfFemale}_{j,t-1}$ is the proportion of females in geographical unit g (school and neighborhood) and year t and year t-1 respectively. The basic idea here is to compare the collective decisions of students (to enrol in college and migrate to another city to pursue tertiary education) from consecutive years who have similar characteristics but the percentage of female peers varies from one year to another. Using the proportion of girls in someone's last year's reference group as an IV relies on the assumption that this proportion has no other effect on someone's decision to enrol in college than through its effect on last year's mean college enrollment and thus this year's someone decision to enrol in college.

The second stage regressions are as follows:

$$IfEnrolled_{igt} = \delta 1 + \kappa_1 \overline{IfEnrolled_{-i,g,\mathbf{t}}} + \psi_1 X_{i,g,t} + T_t + S_g + \lambda_{1g} year + \epsilon_{1ist} (4)$$

$$IfEnrolled_{igt} = \delta 2 + \kappa_2 \overline{IfEnrolled_{-i,g,\mathbf{t-1}}} + \psi_2 X_{i,g,t} + T_t + S_g + \lambda_{2g} year + \epsilon_{2ist} (5)$$

$$g \in \{\{school\}, \{neighborhood\}\}$$

Our key identifying assumption requires that changes in the proportion of female peers within a school and within a neighborhood are not correlated with changes in unobserved factors that could affect students' decisions. In particular, it is required that changes in the proportion of females within schools and neighbourhoods are not associated with changes in student characteristics ie. age, ethnicity income, parental education. We provide evidence that these changes in the proportion of girls within a school and within a neighbourhood are not correlated with changes in school enrolment.

Notice that we exploit within school and within neighbourhood variation from one cohort to another. Our analysis does not look at differences in the percentage of females across schools or neighbourhoods. Additionally, we look at the effect of one's peers on their decision to enrol in college and migrate to another city. To do this, we control for one's performance in the senior year national standardised exams.

The fact that students are assigned to schools based on distance alleviates the concern that students respond to these random shocks in gender composition by switching to another school. Students need to provide adequate evidence of residence in a given region in order to have access to the closest in terms of distance school. But even if students could switch schools, then it would be very difficult to choose the destination school based on the percentage of girls in this school for the following reason: the average percentage of female peers by school or neighbourhoods is not publicly known. But even if it was known it would be difficult to know the percentage of females for a cohort that enters the school in a specific year. Additionally, we provide evidence that leaving a school (drop out or transfer) is not be correlated to the percentage of female students in that school. It is important to note that any factor affecting the proportion of girls in all geographic units in the same way, such as a female fertility decline 17 years before, will be captured by year fixed effects and would thus not invalidate the identification strategy. We include school or neighbourhood fixed effects to control for school or neighbourhood-invariant unobserved factors respectively. One could be worried that time-varying factors ie. better teachers in some years or a new college in the neighbourhood could affect mean enrolment. To address this concern, we include school- or neighbourhood-specific time trends to control for time-varying factors that could be correlated with changes in the fraction of enrolled students in one's reference group.

Next, we turn to academic mobility. We believe that there might exist social interaction effects in the decision to migrate. We model a person's decision to move to a different city in order to pursue tertiary education, given that they were admitted to some college. This decision is a function of the average decision in one's environment as specified in our regression model:

$$IfMigrate_{igt} = \alpha_1 + \gamma_1 \overline{IfMigrate_{j,g,t}} + \beta_1 X_{i,g,t} + \kappa_1 T_t + \mu_1 S_s + \phi_{1g} year + \epsilon_{1ist} (6)$$

$$IfMigrate_{igt} = \alpha_2 + \gamma_2 \overline{IfMigrate_{j,g,t-1}} + \beta_2 X_{i,g,t} + \kappa_2 T_t + \mu_2 S_s + \phi_{2g} year + \epsilon_{2ist} (7)$$

where $IfMigrate_{igt}$ is the decision of student i in geographical unit g and year t and t-1 respectively to migrate in a different city in order to study, conditional on being accepted to college. $\overline{IfMigrate_{-i,g,t}}$ and $\overline{IfMigrate_{j,g,t-1}}$ are the fractions of students except of student i who migrated to a different city in order to study in geographical unit g and year t and t-1 respectively. We include year fixed effects in order to control for time-invariant unobserved characteristics that could affect the migration decision. When we exploit the within school cohort-to-cohort change in the percentage of female students, we include school fixed effects. When we do the analysis at the neighborhood level, we include neighborhood fixed effects and we exploit the differences in school characteristics in a given year within each neighborhood.

We use an instrumental variable approach in order to estimate the effect of social interaction on the decision of students to move to another city to attend college. Again gender composition seems a likely candidate for an instrumental variable. The proportion of females in a geographical unit g may create an environment more conducive to collective migration as exhibited by average patterns of behavior but it has no direct effect on an individual's decision to migrate.

The first stage regressions is as follows:

$$\overline{IfMigrate_{g,t}} = \phi + \kappa \overline{IfFemale_{g,t}} + \beta X_{i,g,t} + \kappa T_t + \mu S_g + e_{g,t} (8)$$

$$\overline{IfMigrate_{g,t}} = \phi + \kappa \overline{IfFemale_{g,t-1}} + \beta X_{i,g,t} + \kappa T_t + \mu S_g + e_{g,t} (9)$$

$$g \in \{\{school\}, \{neighborhood\}\}$$

Our main specifications are estimated at the neighborhood level. When estimated at the geographical units of neighborhood, these specifications address both the endogeneity and simultaneity issues.

Potential threats to our analysis may include the following: Actual networks may be very different from ecologies in one's vicinity. In addition, social media may allow for peer effects that are independent of proximity and render our analysis of spatial social interactions irrelevant. This is less of a fear though as internet penetration is relatively low in Greece[13]. Parents, relatives and much older individuals in a student's environment may influence his/her academic decisions more than his/her same-cohort or one year older peers within his school and/or within his neighborhood .

3.4 Validity of Identification Strategy

Our identification strategy requires that fluctuations in the proportion of female students within a school and within a neighborhood should not be correlated with other cohort-to-cohort changes that could affect students' education decisions. In particular, we check if changes in the proportion of female students within a school and within a neighborhood are correlated with changes in students' observable characteristics. For the universe of students (N=355,808 students) the only characteristics we know are: the age of students and if a student enrolled early in school. This is the case if a student is born in the first quarter of his birth year.

However for a smaller sample of 45 schools (observations=18,670) we also know the ethnicity of students. In Table 3.4, we present some evidence that the schools in the smaller sample have no different characteristics compared to the whole population. We cannot implement the whole analysis based on this smaller sample because we need the universe of students and schools in order to construct the neighbourhoods and exploit within neighbourhood variation. We use this smaller sample of schools to check if changes in the proportion of girls are correlated with changes in students' ethnicity and mobility rates.

Tables 3.2 and 3.3 provide evidence on the balancing tests for the whole sample and the sub-sample of the 45 schools. Table 3.2 reports the estimated coefficients from the OLS regression and a within school regression (school fixed effects) of students' characteristics on the proportion of females in each school. We also report the estimated coefficients from a within school regression when school specific time trends are added (columns (3) and (6)). Table 3.3 reports the estimated coefficients from the within neighbourhood regression (neighborhood fixed effects) with

[13] This is more understandable when one takes into account that Greece has 227 inhabited islands, most of which are quite far from the mainland and have outdated telecommunications infrastructure (Ellinikos Organismos Tourismou (EOT), "Greek islands", April 2012).

(columns (3) and (6)) and without (columns (2) and (4)) adding neighbourhood linear time trends. Again the OLS estimates are reported as a point of comparison.

As we notice from these two tables, the proportion of females is not related to most of the students' characteristics, both in the OLS and the within school/ neighborhood regressions. There are some exceptions in the OLS and within school regression. In particular, the proportion of females within a school seems to be negatively correlated with the proportion of students with Polish and Bulgarian origin, however these correlations are reduced and become statistically insignificant when we add school linear time trends. Within neighbourhoods we find no association between the proportion of females within a neighborhood and students' observable characteristics. All the regressions include year fixed effects. These results suggest that cohort-to-cohort changes in the proportion of female students within a school and within a neighborhood seem to be uncorrelated with changes in students' observed characteristics.

We also examine whether changes in the proportion of female students within a school and within a neighborhood are related to changes in the logarithm of school enrollment. As reported in the first row of Table 3.2 there seem to be a negative association between changes in the proportion of females within a school and changes in the logarithm of school enrolment. Both, the OLS and within school regressions produce estimates negative and statistically significant at 10% . However, this correlation largely reduces and becomes insignificant when school specific time trends are added.

One could still have concerns that students might react to the unpredicted changes in gender compositions. Although students are assigned to schools based on geographical characteristics and it is not easy to switch school, one could still be worried that students might drop out from or switch to another school after being exposed to this information. For example, students who are in schools where the proportion of girls is high/low could drop out. Or transfers of students could be observed that might be correlated to the observed proportion of females in a given school. We address this concern by looking at the correlation between the proportion of female students in a school and the probability that a student drops out from or switch to another school in that year. We use the smaller sample of schools because only for these schools we have data for multiple years and we can identify students who drop out and transfers.

Our dependent variables are: a dummy variable that takes the value of one if the student drops out from school and a dummy that takes the value of one if the student is transferred to this school at the beginning of the school year. Table

3.5 reports the outcome means and the regression estimates separately for boys and girls. The first row in each panel indicates that students' mobility from and to a school is low. Approximately 8% of boys and girls drop out from school in the twelfth grade and around 6-8% of boys and girls respectively transfer to another school at the beginning of the twelfth grade. The second row in each panel reports the regression estimates when school linear trends as well as school and time fixed effects are added. All estimates are small and statistically insignificant. Overall, changes in the proportion of females within a school seem to be uncorrelated with students' mobility across schools and drop out rates.

3.5 Results and Discussion

Table 3.6 shows the linear probability model estimates for the decision to enrol in college. Columns (1) and (2) report the effects of the proportion of enrolled students in year t on a student's decision to enrol in college in year t. Columns (3) and (4) report the effects of the proportion of enrolled students in year t-1 on a student's decision to enrol in college in year t. Each cell in the first and second row in Table 365 shows the estimated coefficient from a separate regression. The estimates presented are based on four different specifications. All specifications include track and year fixed effects. Columns (1) and (3) include school fixed effects and school specific linear trends. Columns (2) and (4) include neighbourhood and neighbourhood specific linear trends. In all specifications we control for a student's gender and admission score. We also include a dummy for students who were born in the first quarter of each year, following Angrist and Krueger 1992, who found significant differences in school outcomes for those students.

The coefficients of interest are positive in year t and statistically significant, revealing strong positive externalities at all levels. An increase of a hundred percent in the proportion of same-age school peers who enrol in college increases one's probability of enrol in college by 8.6 percent, ceteris paribus. This effect decreases at the neighborhood level. In particular, an increase of a hundred percent in the proportion of same-age neighbours who enrol in college increases one's probability of enrol in college by 4.3 percent, ceteris paribus. Coefficients of interest are negative for year t-1 and not very precise.

However, OLS estimates are likely to be bias due to endogeneity issues and the reflection problem. To address these but also further potential unobserved heterogeneity issues, we employ the novel identification strategy of relying on variation in gender composition to explain differences in mean college enrollment in school and neighborhood level. We use an instrumental variable approach to explore social interactions in space and time. Our instrument, gender composition, is likely to affect mean college enrollment since female-heavy school environments are found to be less disruptive and less violent (Lavy and Schlosser 2011).

Tables 3.7 and 3.8 report first and second stage estimates, respectively. Both tables distinguish between social interactions among same-age peers and one-cohort-older peers. Each cell in the first and second row in Table 3.7 shows the estimated coefficient from a separate regression. In our setup, the proportion of girls is a strong predictor of mean enrollment as all first stage estimates are positive and statistically significant at 1%. As we observe in Table 3.7, our instrument is a better predictor of mean enrollment at the school rather than the neighborhood level. In particular, an increase of a hundred percent in the proportion of same-age girls within a school increases mean enrollment by 13.3 % whereas an increase of a hundred percent in the proportion of same-age girls within a neighborhood increases mean enrollment by 8.5%. When we consider last year's proportion of girls then the coefficient of interest declines. In particular, a 100% increase in the percentage of girls in the previous cohort within a school increases this year's mean enrollment by 12.3 %. Furthermore, mean college enrollement within a neighborhood increases by 10.2% if the percentage of girls in the previous year increases by 100%.

Our second stage estimates suggest positive social interactions in education decisions through space and time, with the size of the effect depending on the size of the reference group. Each cell in the first and second row in Table 3.8 shows the estimated coefficient from a separate regression. In Table 3.8, we observe that a hundred percent increase in the proportion of students who enrol in college within one's school in a given year, increases a student's probability to enrol in college by 12.6 % in the same year. Similarly, a student is 7% more likely to enrol in college in a given year if the proportion of students who enrol in college in his neighborhood increases by a hundred percent in that year. We find positive and significant spillover effects among peers in consecutive cohorts. Intuitively, social interactions among students of consecutive cohorts are important, as older peers may function as role models or may provide access to information. We find that a hundred percent increase in the proportion of students attending college within one's school or within one's neighborhood a year before, increases his probability of enrolling in college by 29.1% or 9.6 % percent respectively. Year and track fixed effects are included in all specification. When we exploit within school variation, we control for school fixed effects and school specific time trends. When we use within neighborhood variation, we control for neighborhood and neighborhood specific time trends.

Moreover, we explore social interactions in the decision to study in a different city. Educational mobility is found in the literature to be greatly affected by social norms, labor market structure and income (Tremblay 2005). We focus on those students who enrol in college between 2004 and 2009 (sample size: 355,808). Our models include controls for school or neighbourhood, year and area unobserved time-invariant characteristics. We begin our analysis by estimating specifications (6) and (7) using standard OLS. Our estimates reveal positive social interactions among same-cohort peers and smaller positive externalities coming from students in the previous cohort. Table 3.9 reports the effects of the proportions of migrated students on the decision to migrate of same-cohort or one-cohort-older students using a linear probability model. These estimates show mostly a negative relationship between mean migration and a student's decision to migrate to another city. However, the linear probability model estimates are biased due to reflection and endogeneity. Thus we use the proportion of female peers in one's reference group as an instrumental variable.

Table 3.10 reports first stage estimates. Each column is coming from a separate regression. All coefficients of interest are positive and statistically significant. Again the percentage of female students is a better predictor for the mean migration within a school rather than within a neighbourhood. Ou first stage estimates suggest that changes in the percentage of female peers have significant effects on mean migration in school and neighbourhood among same-cohort students but also in consecutive cohorts. The estimates in columns (1) and (2) are higher than the estimates in columns (3 and (4) respectively implying that the effects are stronger in year t rather than t-1.

Our second stage estimates are reported in Table 3.10. Each column in based on a separate regression. The coefficients of interest are all positive. Our findings suggest significant positive externalities among same-cohort students but significant and smaller positive externalities among students in consecutive cohorts.

3.6 Conclusion

In this paper we have estimated the effects of social interactions on a student's education decisions of college enrollment and academic mobility. Despite the vast literature on the topic, two crucial identification challenges remain: common correlated group effects and simultaneity.

Our contribution to the literature is twofold. First, we propose a new approach in alleviating challenges in identifying spillover effects by using time lagged group characteristics. Second, we provide evidence on social interactions using a special institutional setting that allows for spatial variation of group characteristics. So far, the existing literature on social interactions has focused almost exclusively on scholastic performance. The only exemptions to our knowledge are Sacerdote 2011 who identify the effect of social interactions on drinking, drug use, and criminal behavior and Giorgi et al. 2007 who find significant effects on the choice of college major.

When social interactions are not taken into account, educational treatments may result in misallocation of resources and may fall short of policy goals. Our results aim to inform public policies that target ability mismatch.

We employ instrumental variable techniques to estimate utility linkages at different space and time levels. We battle the reflection problem and the endogeneity issues by using time lagged school and neighborhood student gender compostion as an instrument. Using repeated cross-sectional data, we exploit within-school and within-neighborhood cohort-to-cohort variation to examine the effect of random changes in gender composition on mean college enrollment. Then we look at the effect on a student's decision to enrol in college.

We find that the choices of a student's peers affect their decision to enrol in college and migrate to another city to pursue tertiary education. We use a novel dataset from Greece that contains the universe of high school graduates from 2004 to 2009. We focus our analysis on four reference groups: same-cohort peers in school, one-cohort-older peers in school, same-cohort peers in neighbourhood and one-cohort-older peers in neighbourhood.

Our evidence supports the hypothesis that individuals derive utility from conformity or have access to information, with the size of the externality decreasing in space distance. Our results show that one is more likely to enrol in college and move to another city to pursue post secondary education when many of his peers make the same choices. A hundred percent increase in the percentage of one-cohort-older peers within a school and within a neighborhood who enrolled in college increases a student's probability of college enrollment by 29.1 and 9.6 percent, respectively. In addition, a hundred percent increase in the percentage of same-cohort students who enrol in college within a school and within a neighborhood increases one's own probability to enrol in college by 12.6 and 7 percent respectively.

While our paper has examined several important determinants of college enrollment and migration decision, several avenues of future research remain. Understanding the mechanism that underlies social interactions is the next big question in the literature. Future research could push forward the front of understanding the mechanism that underlies social interactions.

Table 3.1: Descriptive Statistics

	Mean	**Std. Dev.**	**Min.**	**Max.**	**N**
Panel A: Individual Level					
First quarter of birth	0.16	0.368	0	1	355,808
Female	0.567	0.495	0	1	355,808
National Exams Score	13.16	4.062	0.52	19.95	355,808
If enrolled	0.812	0.391	0	1	355,808
Mobile students	0.748	0.434	0	1	260,472
Specialty in Classics	0.365	0.481	0	1	355,808
Specialty in Natural Science	0.154	0.361	0	1	355,808
Specialty in Technical Studies	0.484	0.5	0	1	355,808
Postcode Income (Euro, 2009)	29,464	8,441	9,573	122,879	355,808
Aggregate Enrollment	60,206	6,372	52,450	68,136	355,808
Panel B: School Level					
Private	0.081	0.266	0	1	1,319
Income if private (Euro, 2009)	30,575	18,378	16,085	122,879	1,319
National score if private	13.69	2.70	4.7	17.34	1,319
Experimental	0.022	0.149	0	1	1,319
Income if experimental (Euro, 2009)	29,754	14,775	17,583	74,798	1,319
National score if experimental	14.40	1.00	12.23	16.17	1,319
Public	0.89	0.31	0	1	1,319
Income if public (Euro, 2009)	19,327	5,565	9,573	74,798	1,319
National score if public	12.26	1.56	2.97	16.36	1,319
Urban	0.898	0.301	0	1	1,319
Distance to nearest college campus(in miles)	10.871	24.083	0.105	1095.452	1,319
No of students in each school	46	34	0.16	179	1,319
Panel C: Neighborhood Level					
No of schools in each neighborhood	4.449	5.014	2	35	250
No of students in each neighborhood	929.291	1,246.298	8	10,559	250

Note: Data span six cohorts 2004-2009 of 60.119 students on average. Number of schools: 1319. Among those 413 high schools are in Athens or the surrounding suburbs. The national exam score ranges from 0 to 20. Mobile students are those who move to a different city in order to study.

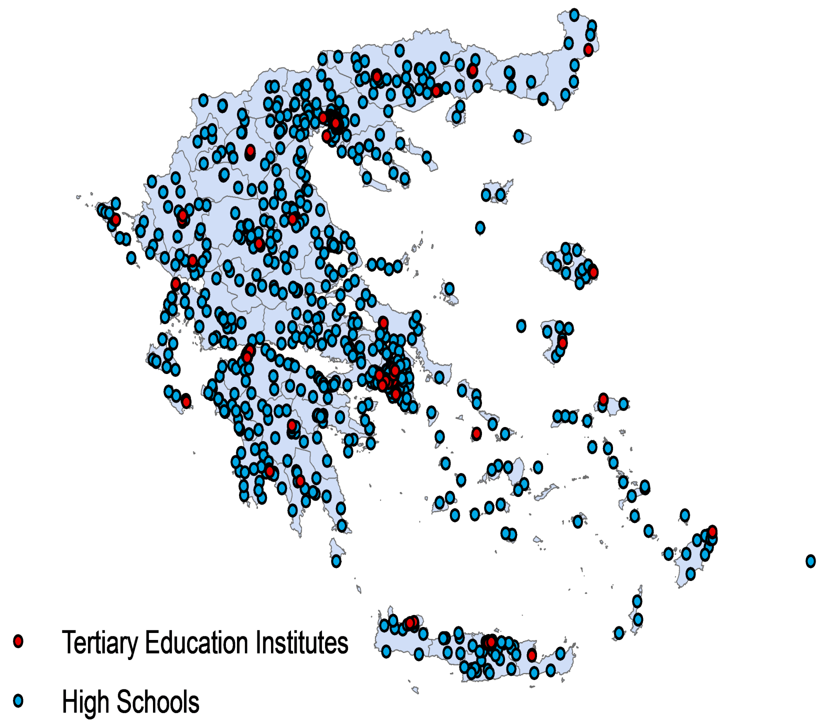


Figure 3.1: Map of schools

Table 3.2: BALANCING TESTS FOR PROP. OF FEMALES IN SCHOOL

	WHOLE SAMPLE			SMALLER SAMPLE		
	OLS	School FE	School FE +school linear time trends	OLS	School FE	School FE +school linear time trends
	(1)	(2)	(3)	(4)	(5)	(6)
logEnrollment	-0.075	-0.074	0.004	-0.125	-0.125	0.006
	(0.036)*	(0.036)*	(0.024)	(0.053)*	(0.054)*	(0.048)
EarlyEnrollment	-0.002	0.005	0.0009	-0.003	-0.006	0.0010
	(0.019)	(0.019)	(0.001)	(0.016)	(0.016)	(0.001)
Age	-0.002	0.002	-0.002	-0.003	-0.001	-0.004
	(0.002)	(0.002)	(0.003)	(0.010)	(0.010)	(0.006)
Ethnicity						
Greece				0.001	0.001	0.003
				(0.001)	(0.001)	(0.002)
Albany				0.004	0.004	0.016
				(0.013)	(0.013)	(0.011)
Bulgaria				-0.060	-0.060	-0.033
				(0.027)*	(0.027)*	(0.022)
Italy				-0.010	-0.010	-0.001
				(0.011)	(0.011)	(0.006)
Russia				-0.012	-0.012	0.011
				(0.027)	(0.027)	(0.017)
Poland				-0.054	-0.054	-0.022
				(0.024)*	(0.024)*	(0.019)
Ukraine				-0.003	-0.003	0.004
				(0.016)	(0.016)	(0.013)
N	355,808	355,808	355,808	18,670	18,670	18,670

Note: Standard errors are clustered at the school level. A constant is also included. *,**,*** denotes significance at the 10%,5% and 1% level respectively. The table reports OLS and school fixed effects estimates from separate regressions. Columns (3) and (6) report school fixed effects estimates having added school linear time trends. Year dummies are included in all regressions.

Table 3.3: BALANCING TESTS FOR PROP. OF FEMALES IN NEIGHBOURHOOD

	WHOLE SAMPLE			**SMALLER SAMPLE**		
	OLS	neighb. FE	neighb. FE+ neighb. linear time trends	OLS	neighb.FE	neighb. FE+ neighb. linear time trends
	(1)	(2)	(3)	(4)	(5)	(6)
logEnrollment	-0.081	-0.080	0.002	-0.089	-0.089	0.003
	(0.048)	(0.050)	(0.035)	(0.050)	(0.055)	(0.031)
EarlyEnrollment	0.002	0.005	0.001	0.004	0.004	0.001
	(0.003)	(0.003)	(0.035)	(0.007)	(0.007)	(0.023)
Age	0.001	0.003	-0.002	0.003	0.003	-0.001
	(0.001)	(0.001)	(0.005)	(0.004)	(0.003)	(0.002)
Ethnicity						
Greece				0.001	0.001	0.003
				(0.001)	(0.001)	(0.002)
Albany				-0.007	-0.007	-0.004
				(0.006)	(0.006)	(0.003)
Bulgaria				0.002	0.002	0.008
				(0.009)	(0.009)	(0.006)
Italy				-0.004	-0.004	-0.002
				(0.004)	(0.004)	(0.002)
Russia				0.005	0.005	0.009
				(0.009)	(0.009)	(0.006)
Poland				-0.002	-0.002	0.005
				(0.010)	(0.010)	(0.004)
Ukraine				-0.010	-0.010	-0.009
				(0.006)	(0.006)	(0.005)
N	355,808	355,808	355,808	18,670	18,670	18,670

Note: Standard errors are clustered at the neighborhood level. A constant is also included. *,**,*** denotes significance at the 10%,5% and 1% level respectively. The table reports OLS and neighborhood fixed effects estimates from separate regressions. Columns (3) and (6) report neighborhood fixed effects estimates having added neighborhood linear time trends. Year dummies are included in all regressions.

Table 3.4: Descriptive statistics for smaller sample and population

	Smaller Sample Mean	**Population Mean**	**Difference**	**Std. Dev.**
log Postcode income	9.962	9.968	0.006	(0.014)
Private school	0.080	0.081	-0.001	(0.001)
Public school	0.897	0.899	-0.002	(0.003)
Experimental school	0.020	0.022	0.002	(0.003)
Urban	0.899	0.898	0.001	(0.001)

Note: 18,670 obs. in smaller sample and 355,808 obs. in population. 45 schools in sample, 1319 schools in population.

Table 3.5: Estimation results : Drop out and Transfers

Dependent Variable: Dummy for drop out and Transfers		
	(1)	(2)
Variable	(Males)	(Females)
Drop out		
Outcome mean	0.080	0.078
Regression estimates	0.060	0.020
	(0.046)	(0.038)
Transfers		
Outcome mean	0.068	0.075
Regression estimates	0.020	-0.052
	(0.075)	(0.071)

Note: The table reports means of the dependent variable (first row) and estimates (second row) for the effects of the proportion of females on the probability that a student leaves school the following year. We use the smaller sample here of 45 schools. Clusters at school level. All regressions include controls for student characteristics. Standard errors are clustered at the school level. All regressions include school fixed effects, year fixed effects and school linear time trends.

Table 3.6: Linear Probability Model Estimates

	Dependent Variable: College Enrollment			
	(1)	(2)	(3)	(4)
Sample:	School	Neighborhood	School	Neighborhood
% Enrolled$_t$	0.086	0.043		
	(0.011)***	(0.012)***		
% Enrolled$_{t-1}$			-0.039	-0.017
			(0.012)***	(0.014)
Born in 1st quarter	0.006	0.006	0.007	0.007
	(0.001)***	(0.001)***	(0.001)***	(0.001)***
Female	-0.003	-0.003	-0.003	-0.003
	(0.001)**	(0.001)***	(0.001)**	(0.001)***
Admission Score	0.069	0.069	0.069	0.069
	(0.000)***	(0.000)***	(0.000)***	(0.000)***
Speciality FE	✓	✓	✓	✓
Year FE	✓	✓	✓	✓
School FE	✓	✓	✓	✓
School specific time trends	✓		✓	
Neighbourhood specific time trends		✓		✓
N	355,808	355,808	355,808	355,808
R^2	0.47	0.47	0.47	0.47

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Standard errors are clustered at the school level. An intercept is also included.

Table 3.7: First stage estimates

Dependent variable:	**Mean College Enrollment$_t$**			
	school	neighborhood	school	neighbourhood
Proportion of girls$_t$	0.133 (0.019)***	0.085 (0.030)***		
Proportion of girls$_{t-1}$			0.123 (0.018)***	0.102 (0.003)***
Female	-0.0004 (0.0003)	-0.0008 (0.0002)***	-0.0001 (0.0003)	-0.0004 (0.0003)
Admission Score	0.001 (0.0008)***	0.0002 (0.00004)***	0.00002 (0.0003)***	0.00003 (0.00003)
Born in first quarter	0.0001 (0.00004)	-0.0006 (0.0002)***	0.0002 (0.0002)	0.0008 (0.0007)
Speciality FE	✓	✓	✓	✓
Year FE	✓	✓	✓	✓
School FE	✓	✓	✓	✓
School specific time trends	✓		✓	
Neighbourhood specific time trends		✓		✓
N	355,808	355,808	355,808	355,808
R^2	0.54	0.70	0.43	0.62
F-statistic 1st stage	14.22	17.57	12.11	18.13

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Standard errors are clustered at the school level. An intercept is also included.

Table 3.8: IV Second Stage Estimates

	Dependent Variable: College Enrollment			
	(1)	(2)	(3)	(4)
Sample:	School	Neighborhood	School	Neighborhood
% Enrolled$_t$	0.126	0.070		
	(0.036)***	(0.033)**		
% Enrolled$_{t-1}$			0.291	0.096
			(0.037)***	(0.035)***
Admission Score	0.069	0.069	0.069	0.069
	(0.001)***	(0.000)***	(0.000)***	(0.000)***
Female	-0.003	-0.003	-0.003	-0.003
	(0.001)***	(0.001)***	(0.001)***	(0.001)***
Born in first quarter	0.006	0.006	0.006	0.007
	(0.001)***	(0.001)***	(0.001)***	(0.001)***
Speciality FE	✓	✓	✓	✓
Year FE	✓	✓	✓	✓
School FE	✓	✓	✓	✓
School specific time trends	✓		✓	
Neighbourhood specific time trends		✓		✓
N	355,808	355,808	355,808	355,808

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Standard errors are clustered at the school level. An intercept is also included.

Table 3.9: LPM Migration Decision

	Dependent Variable: Migration Decision			
	(1)	(2)	(3)	(4)
Sample:	School	Neighborhood	School	Neighborhood
% Migrated$_t$	-0.099	0.095		
	(0.035)***	(0.037)***		
% Migrated$_{t-1}$			-0.087	-0.067
			(0.026)***	(0.022)***
Born in 1st quarter	-0.001	-0.001	-0.001	-0.001
	(0.002)	(0.002)	(0.002)	(0.002)
Female	-0.013	-0.013	-0.013	-0.013
	(0.002)***	(0.002)***	(0.002)***	(0.002)***
Admission Score	-0.023	-0.023	-0.023	-0.023
	(0.002)***	(0.002)***	(0.002)***	(0.002)***
Speciality FE	✓	✓	✓	✓
Year FE	✓	✓	✓	✓
School FE	✓	✓	✓	✓
School specific time trends	✓		✓	
Neighbourhood specific time trends		✓		✓
R^2	0.30	0.30	0.30	0.30
N	355,808	355,808	355,808	355,808

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Standard errors are clustered at the school level. A intercept is also included.

Table 3.10: First stage estimates for migration decision

	Dependent Variable: Mean Migration Decision			
	(1)	(2)	(3)	(4)
Sample:	School	Neighborhood	School	Neighborhood
Proportion of girls_t	0.125	0.098		
	(0.023)***	(0.020)***		
Proportion of girls_{t-1}			0.114	0.089
			(0.027)***	(0.022)***
Admission Score	0.0002	0.001	0.0006	0.001
	(0.0001)*	(0.0001)***	(0.0003)***	(0.0001)***
Female	0.002	0.003	0.001	0.002
	(0.0008)**	(0.007)*** (0.0008)	(0.0007)***	
Born in 1t quarter	-0.0006	-0.0007	-0.006	-0.0008
	(0.001)	(0.001)	(0.001)	(0.001)
Speciality FE	✓	✓	✓	✓
Year FE	✓	✓	✓	✓
School FE	✓	✓	✓	✓
School specific time trends	✓		✓	
Neighbourhood specific time trends		✓		✓
N	260,472	260,472	260,472	260,472
R^2	0.37	0.39	0.37	0.39
$F - statistic1ststage$	16.20	14.8	15.9	14.7

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Standard errors are clustered at the school level. An intercept is also included.

Table 3.11: IV Estimates for Migration Decision

	Dependent Variable: Migration Decision			
	(1)	(2)	(3)	(4)
Sample:	School	Neighborhood	School	Neighborhood
% Migrated$_t$	0.107	0.091		
	(0.032)***	(0.027)***		
% Migrated$_{t-1}$			0.101	0.065
			(0.036)**	(0.034)*
Admission Score	-0.022	-0.023	-0.023	-0.023
	(0.000)***	(0.000)***	(0.000)***	(0.000)***
Female	-0.014	-0.014	-0.014	-0.014
	(0.002)***	(0.002)***	(0.002)***	(0.002)***
Born in 1st quarter	-0.001	-0.001	-0.001	-0.002
	(0.002)	(0.002)	(0.002)	(0.002)
Speciality FE	✓	✓	✓	✓
Year FE	✓	✓	✓	✓
School FE	✓	✓	✓	✓
School specific time trends	✓		✓	
Neighbourhood specific time trends		✓		✓
N	260,472	260,472	260,472	260,472

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Standard errors are clustered at the school level. An intercept is also included.

Chapter 4

Externalities in the Classroom: Identification of returns to absences and peer effects

4.1 Introduction

Most educational systems rely on lectures and class meetings as a means of instruction. This is even more prevalent when secondary or pre-tertiary education is considered. Nevertheless, class attendance is not always perfect. Lecture learning is based on group learning, which may not be the optimal learning style for everyone. As a result, many students decide to skip class when given the opportunity. In a classroom, students compete for the attention and time of the instructor. Thus, their consumption of education induces externalities on one another. Romer [1993] claims that college students in three elite U.S. universities were found to perform better when attending classes and completing homework. Nevertheless, this claim may apply for only a small part in the right tail of the ability distribution in a given society. Lectures in classrooms with samples that reflect the actual ability distribution of students may not run completely smoothly. To give an example, students who act up or disrupt the lecture may be more likely to be found in non-elite schools. The question that arises here is whether someone should attend class or stay and study at home given their ability ceteris paribus.

The literature regarding class absenteeism is divided into two main categories: one refers to the reasons for students being absent from class (Levine 1992, Chong et al. 2009) and the second one is concerned with the effect of students' absenteeism on their scholastic outcomes (Romer 1993, Caviglia Harris 2006, Chen and Lin 2008, Arulampalam et al. 2012, Latif and Miles 2013). Most of these papers use college and field specific class attendance data. In particular, most of these papers use data regarding Economics, Accounting or Management students. The majority of these studies look at correlations and find a negative relationship between students' absenteeism and academic performance or a negligible one (Caviglia Harris 2006). Evidence from the existing literature suggests that class attendance improves educational outcomes. Lin and Chen 2006 using a sample of 129 college students in Taiwan find a 4% exam score improvement associated with higher class attendance. A subsequent study by the same authors Chen and Lin 2008 involved an experiment where different sections of the same college course were subject to random changes in the curriculum although everyone sat the same exam at the end of the semester. The authors found that having the instructor cover all of the material improved score by as high as 18%. Latif and Miles 2013 used panel data of exam scores of Canadian college students to measure the effect of class attendance on exam performance. They find that when controlling for student heterogeneity, exam performance is positively related to class attendance. This is the first study that attempts to control for endogeneity but no exogenous variation is exploited. Similar results have been obtained when college classes on science Moore 2006 or economics Cohn and Johnson 2006 are considered. Arulampalam et al. 2012 use panel data to identify the causal relationship between class attendance and students' University performance. This is the only causal study in this literature. Focusing on Economics students, they use quantile regression analysis and find that skipping classes leads to poorer performance. Interestingly, they highlight that the relationship between class attendance and students' performance may vary with student ability. Caviglia Harris 2006 examines the impact of mandatory attendance of microeconomic classes on students' college performance. After accounting for students' motivation, he finds that class attendance does not impact grades. This is the only paper that finds a negligible effect between class attendance and students' academic outcomes. Despite the rich literature that involves college data, there is little evidence that the same results hold in a less filtered context, like high schools.

In this paper, we investigate the causal relationship between class attendance and exam performance. Our approach exploits a natural experiment that increased the absence allowance of high school students by fifty hours only if their grade point average exceeded a threshold in the previous grade. In our context, senior year students are maximizing their end-of-year test scores by choosing how much time to spend in and outside classroom. The end-of-year exam performance is very crucial for students' post-secondary placement because it determines the university entrance score. The treatment offers exogenous variation by relaxing the budget constraint only for some students, whose marginal utility of time may be higher than the average.

In the institutional setting examined here high school students in the senior year usually prepare for the university admission exams. Admission to tertiary education is based solely on test scores achieved at the end of the senior year. In order to apply for university admission, students take exams in a specific number of subjects once per year. In this context, students are to allocate studying time between attending classes in school and studying at home for the national exams. The basic idea behind the reform was to relax the budget constraint for the high-achieving students and allow them to strategically manipulate the additional hours of excused absences they were provided with. In the second term and just before the national exams, it might be beneficial for high achieving students to spend time at home revising past exams or focusing on specific subjects than attending classes in school.

Using an instrumental variable method, we identify the causal effect of class attendance on exam performance. We exploit exogenous variation generated by the reform to examine the causal path between attendance and school performance in a natural experiment context. We control for individual-specific, school-specific, year-specific and grade-specific heterogeneity by using longitudinal data on exam performance of students in consecutive grades.

Our findings suggest that even when students can choose when to be absent from class, the returns to absences are negative. It is not very common for students to be allowed to choose when to skip classes. We find that students are more absent from class by 15 hours on average when there are provided with the additional hours of excused absences. We also find that an additional day of absence from school decreases a student's gpa and math score by 0.1 and greek language score by 0.08 of a standard deviation. The effects are larger for female students, indicating that female students may lose more from skipping classes. Our results suggest that attendance is an important driver of school performance.

To our knowledge, this is the first paper that identifies the returns to absences using a quasi-experimental approach. By using data of students who attend public schools, which is the case for more than ninety percent of the universe of high school students, we avoid truncating the observed support of the ability distribution. We have collected transcript data of the three last grades of high school from 98 schools in Greece. The lack of selection issues allows us to identify returns to absences which contributes to the external validity of our study.

The remaining of the paper is organised as follows: Section 4.2 describes a description of the institutional setting. Section 4.3 describes the data. Section 4.4 presents a regression discontinuity approach and section 4.5 proposes an identification using an absences law instrument. Section 4.6 presents some heterogeneous effects and lastly section 4.7 concludes.

4.2 Background

It is useful to provide some background on the design of the institutional setting in which our natural experiment takes place. Public high schools are the norm in Greece as only around 8 percent of students attend private high school[1]. Assignment to high school schools is based on geographical proximity, namely a school district system. Every high school offers the same curriculum and funding is a linear function

[1] Descriptive statistics from a dataset that covers the universe of high school graduates-dataset used in the previous Chapter of this thesis- between 2003 and 2011 show that 90% of students attend public schools, 2% attend public experimental (charter) schools and 8% attend private high schools. There are 1319 high schools in Greece, of which 112 are private and 23 are experimental.

of the number of students. Teachers' quality characterictics such as education and experience are not taken into account for allocation of teachers to schools. By law, assignment to classrooms is based on alphabetical order.

Up until the end of the school year 2005-2006, every student could have 50 hours of unexcused and 64 hours of excused absence from class within a given year. So a student can be absent from class fo 114 hours in total without a penalty. An hour of absence can be excused only by a doctor or someone with the child's custody -usually the parents. Only whole days of absence can be excused. For example, if a student goes to school late in the morning or if they decide to skip school midday, their absences cannot be excused. The penalty for exceeding the number of allowed absences is to repeat the grade.

In late spring 2006, the Ministry of Education passed a new bill that regulated the number of allowed hours of absence from school. The new policy was not pre-announced or discussed publicly before it was announced. The new bill provided eligible twelfth grade students with 50 additional hours of excused absences. Eligibility was determined based on past grade point average (GPA). In particular, every student who had received a GPA higher than 15/20 in the previous grade (eleventh grade) was eligible to take up more absences this year. In our analysis, we use the graduating class of 2006 as a control group and the graduating class of 2007 as the treated group. We use only the graduating class of 2007 as the treated group because this cohort was surprised by the reform. The timing of the new policy did not allow students to choose a different effort level and manipulate their eleventh grade results. Students had already taken their school end of year exams when the reform was announced.

Students view the number of hours they can be absent from school with no penalty as an "allowance" and they tend to use it right before important exams during the school year in order to prepare. For example, when we look at absences during the first and second semester separately (the school year consists of two semesters), we see that absences in the second semester are much higher, which makes sense as students could better prepare for the final exams when most of the module curriculum is covered. This is even more pronounced for high achieving students. In Figure 4.1, we observe the twelfth grade total absences patterns by term for each decile of tenth grade performance (gpa). High achieving students tend to be more absent from school in the second term while the opposite is observed for low achieving students. It is also unlikely that students are more frequently sick in the second term than in the first one. Weather conditions in Greece would work to the other direction, given that the first school term coincides with the winter period when students are more likely to be sick. This figure provides evidence that high achieving students might strategically manipulate their absences in order to prepare for important exams. The school term is designed in a way that it gives only one study/revision week for preparation after the end of the school term and before the beginning of the exam period.

Students take national-standardised exams that matter for both high school graduation and university admission. The format of the national exams is the same as the one of the school exams in the previous grades and they are externally marked and proctored.

It is worth mentioning that by design lectures of the same subject are usually spread out within the weekly schedule of classes. This is important because one may worry that eligible students might skip classes of a particular subject. This strategic selection of classes is not entirely possible because only whole days of absence can be excused.

Around sixty percent of school subjects are core education ones and the remaining consist of electives and specialization/track courses. Unlike other educational systems, in Greece students remain in their assigned classroom for the majority of school periods instead of moving to different rooms depending on the subject being taught. This setting guarantees that a student's peer group remains the same for a series of courses, including greek language and mathematics, considered in our analysis.

4.3 Data

We have collected hand collected data from a large randomized sample of high schools in Greece. For this study we focus on public schools (Sample: 98 schools, 11,239 students). This novel dataset includes every student that graduated from one of the sampled schools between 2006 and 2007 and contains panel information from the following sources:

1. Administrative data from the High Schools containing course taking information and exam grades in each of the last three years of secondary education, class identifier, class size [2], gender, year of birth and graduation year. For each student we also know how many hours were they absent from their class in the eleventh and the twelfth grade. We know how many absence hours

[2]corr (class size, income)=0.149,corr (class size, experimental)=0.249, corr (class size, urban)=0.179

did the parents excused and how many hours of students' absences remained unexcused.

2. School specific information such as name of school, type of school (private, public[3], experimental[4]), geographical location.
3. The Ministry of Finance provided us with average net income information at the postcode of the school in 2009 Euro.
4. The Ministry of Internal Affairs provided us with urban density information. Urban areas are those with more than 20,000 inhabitants.

Table 4.1 provides summary statistics for our sample in grade 12 (11,239 students). Students are on average 50 hours absent from school in grade 11 and 76 hours absent in grade 12. Out of these, 20 and 42 hours respectively are excused. In addition, 57% of students are females in our sample and 95% of students attend a high school in an urban area.

Table 4.2 presents some mean comparisons between 2006 and 2007 which are the control cohort and the treated cohort respectively. Although we do not observe a significant difference between the percentage of students who are female, attend public/experimental and urban schools, what is actually very interesting is that there is no difference between students' eleventh grade GPA between 2006 and 2007. The eleventh grade GPA determines a student's eligibility status. This confirms our prior belief that the reform came as a surprise to these students.

A limitation of our data is that we do not observe which exact hour a student is absent from class ie. if he skips a class for an important subject or a class for a subject that is not nationally examinable. However, a student cannot choose when to skip some classes within a school day. Otherwise, these hours of absence cannot be excused.

4.4 The effect around the cut-off

4.4.1 Strategy

We start our analysis by looking at students who are around the eleventh grade gpa cut-off (or the eligibility cut-off). By using a Regression Discontinuity design we can identify the effect of the additional absences on students' academic performance in the treated year. So, we will compare students who are just to the left with students

[3] Students are assigned to public schools according to a school district system.
[4] Admission to experimental schools is based on a lottery.

who are just to right of the cut-off value in the treated year (2007). Students with eleventh grade gpa below 15 cannot exploit the additional hours of absences that students with gpa above 15 are offered. Within each school, we rank students according to their eleventh grade gpa and identify those who are above and below the threshold of gpa=15. Let t_0 take the value of 15 and t_i be student's gpa in the eleventh grade.

The first stage regression can be specified as:

$$TA_i = \alpha f_1(t_i) + \psi 1[t_i > t_0] + \omega(1)$$

where $1[t_i > t_0]$ is an indicator for whether a student is eligible to be more absent i.e. if his eleventh grade gpa is greater than or equal to the threshold value of 15/20, t_i is the eleventh grade gpa of student i and $f_1(t_i)$ is a control function for the gpa of student i. We also use specification (1) to apply a first differences approach. We do that because there might be individual specific unobserved characteristics that are omitted and might affect our variables of interest. By taking first differences of outcomes and covariates between twelfth and eleventh grade, we get rid of potential time invariant omitted variables. Specification (1) will tell us if students just to the right of the cut-off use the additional hours of absence, when they are allowed to do so, compared to students who are just to the left of the cut-off and they are not allowed. By using the first difference version of specification (1), we will find how many more hours absences do students to the right of the cut-off use compared to their eleventh grade hours of absences with regards to the counterfactual group of students.

The idea behind the regression discontinuity design which was initially proposed by Thistlethwaite and Campbell [1960] is that discontinuities like the above can be used to identify the causal effect of scoring in the eleventh grade above 15/20. Intuitively, assume that the gpa is smoothly related to characteristics that affect academic performance. Having assumed that, pupils with scores just above the threshold value will provide a proper control group for pupils with scores just below the threshold value. This is visualised in Figure 4.2. Then any differences in the outcomes of those students can be attributed to the fact that some students are eligible to be more absent from school due to the reform.

The reduced form equation that will estimate the effect of being eligible to be more absent from school on academic outcomes, can be described as:

$$Y_i = \delta(t_i) + \gamma 1[t_i > t_0] + e_i(2)$$

where: Y_i is the standardised twelfth grade score in Modern Greek, Mathematics and the twelfth grade gpa for student i ,
Results will be presented for small enough neighborhood areas of different sizes around the cut-off. When we do that, $\delta(t_i)$ will be constant and γ will identify the causal effect of being allowed to be more absent from school on test scores, non-parametrically (Hahn et al. [2001]).

4.4.2 Results

Our results are produced using the non-parametric approach discussed above, implementing a local linear regression constructed with a triangular kernel. The first stage is not very strong. This is something that we can also observe in Figures 4.2 and 4.3. In Table 3, we present the estimates for regression discontinuity estimates using six different bandwidths. In columns 1,2,3 (Table 4.3) we restrict the sample to those students who are 0.5/20 [5], 1/20, 1.5/20 to the left and the right of the cut-off respectively. In columns 4,5,6 (Table 4.3) we increase the bandwidth further using the bandwidths suggested by Calonico et al. [2014], Imbens and Kalyanaraman [2012] and Ludwig and Miller [2007] respectively. Using these bandwidths, we do not find evidence that students just to the right of the cut-off used more hours of allowed absences compared to students just to the left of the cut-off. The coefficients are positive but are statistically insignificant.

Then we present reduced form results (Table 4.4). We present results for each subject separately (modern greek, mathematics and gpa in twelfth grade). We do not find any consistent pattern across columns. Only the first difference estimates in Mathematics (Panel B) are statistically significant in some specifications. Students who are eligible to be more absent from class, experience a decrease in their standardised score in Mathematics by 0.15-0.30 standard deviations.

These results make us think that there is no effect around the threshold but the reform might affect students who are farther away from the cut-off. The impact of class attendance on students' performance may vary with student ability (Arulampalam et al. 2012). The reform might not be used by students who are just

[5] The maximum score a student can get is 20. 0.5/20 corresponds to 2.5 out of 100

to the right of the cut-off. Not all eligible students might be interested in using the additional excused hours of absences that they were provided with. Intuitively, students who are at the top percentiles of the ability distribution might decide to spend more hours preparing for the national exams. The education literature suggests that there is the so called self-regulated learning that is more pronounced for the high achieving students. (Barry J. and Manuel 1990,Nicola and Debra 2006).

The intuition is that high achieving students might be/or believe that they are more constructive when studying at home rather than staying in the classroom, especially if it is noisy. The research on self regulated-learning suggests that these students might set goals for their learning and monitor, regulate, and control their cognition, motivation, and behaviour better than students of lower academic ability. So the marginal utility of spending an additional hour in the class might be different for a student who had an eleventh grade gpa equal to 16 and another student who had a gpa equal to 19 although there are both eligible to use the reform. This makes us think that there might not be an effect around the performance cut-off (15/20).

As shown in Figure 4.2, some covariates exhibit a jump around the threshold that violate the assumptions of identification of the treatment effects using an regression discontinuity approach. The RD approach may be inappropriate for identification if individuals to the left of the cut-off differ in more than one ways from individuals to the right of the cut-off. To control for individual specific drivers of the observed behaviour we take first differences of observed variables between twelfth and eleventh grade. The change in these differences around the cut-off is shown in Figure 4.3.

The regression discontinuity estimates that look at the effect of the reform on school performance seem weak. We suspect that this is due to the fact that it may not be those just above the threshold of eligibility who exploit the new policy and take up more hours of absence but rather those we are in the right tail of the distribution due to self-regulated learning.

4.5 Identification using absences law instruments

4.5.1 Empirical Strategy

We then look at the average effect of the reform on all eligible students and not only those who are just around the cut-off. To do this, we postulate a model where individuals' school performance is a function of own hours of absences:

$$Score_{icsgt} = \alpha_o + \alpha_1 TotalAbsences_{icstg} + \alpha_2 Senior_{it}$$

$$+studentFE + schoolFE + \epsilon_{icst}(3)$$

where the dependent variable $Score_{icsgt}$ is the twelfth grade performance of student i in class c in school s in grade g and in year t. We include a dummy variable $Senior_{it}$ that takes the value 1 if the student is in the twelfth grade to pick up mean changes in effort due to the high stake exams in the twelfth grade. Student fixed effects are used to control for all observed and unobserved student characteristics that are constant over time. This could include student effort and ability, as well as family related factors like help from parents to help their child with the homework. We also include school fixed effects to control for any time-invariant systematic differences across schools.

The dependent variable includes the national exam performance of student i in two different subjects (greek language and mathematics) and the twelfth grade gpa. Notice that we exploit the panel aspect of our data. Standard errors are clustered at the class level. The main coefficient of interest is α_1 indicating how total absences affect a student's performance. Estimating equation 3 using OLS may lead to a number of problems, since the variable "total absences" is endogenous. Bias may be created by omitted variables that may affect the performance of students and their decision to stay at home, such as the existence of modern facilities like interactive boards in some schools. Then the total absences of the student may be correlated with the error term and that would invalidate the OLS estimates. For example, the degree of parental monitoring or other individual characteristics such as self-discipline or the motivation level affect both the hours that students decide to stay at home and student' s productivity. We exclude from our analysis students who enrol into private schools in order to avoid selection issues. Not controlling for unobserved characteristics, would add another estimation bias.

Furthermore, measurement error would bias the estimation of the parameters of interest. The bias from measurement error may be less of a threat when this error is time invariant but even measures of performance and attendance are less than perfect. An instrumental variables approach can address biases due to selection, omitted variables and measurement error. Therefore, we exploit the reform in the absences allowance law to construct an instrument for class attendance.

We mitigate the endogeneity issues by using an instrumental variables approach in order to obtain unbiased estimates of the causal effects of interest. We employ a difference-in-difference approach in order to measure the effect of the reform on total absences. In particular, we interact the eligibility status dummy with the year dummy that takes the value 1 if the new absences law is in place. As outlined before, students with an eleventh grade gpa above 15 out of 20 can be more absent from school in 2007 compared to 2006. Students with an eleventh grade GPA below 15 out of 20 have the same allowance in both years. Thus, we use this reform as a source of exogenous variation for total absences.

The first stage regression is the following:

$$TotalAbsences_{icstg} = \beta_1 + \beta_2 Eligibility_{icst} + \beta_3 Reform_t + \beta_4 Eligibilityicst * Reform_t$$

$$+\beta_5 Senior_{it} + schoolFE + studentFE + \epsilon_{icstg} (4)$$

The main instrument for total absences is the interaction between the eligibility status and the reform year dummy. This interaction term in equation (4) measures the treatment effect of the reform on total absences of eligible students compared to non-eligible ones. The interaction variable is zero for non-eligible students in both years. In 2007, it becomes equal to one only for eligible students. In order for the reform to be an appropriate instrument for total absences, it must be correlated with total absences and only affect national exam performance and twelfth grade gpa through hours of absence.

The outcome variables and the effects of the reform are likely to be correlated for all students in a given class. Thus, we control for any dependence between observations within a class by clustering all results at the class level. Using instrumental variables that stem from the reform relies on the assumption that the reform had no other effect on a student's performance than through its effect on the students absences and peer group quality. It is important to note than any factor coinciding with the reform, affecting all students in Greece in a similar way, such as a possible change in exam difficulty, will be captured by the reform year dummy that takes the value the value 1 for the year 2007 and 0 before. As the unaffected individuals act as a control group, only factors changing at the same time as the reform may be potential threat to our identification strategy. To our knowledge there were no other relevant changes in the institutional setting at the time the reform of interest was implemented.

Lastly, any difference in difference type strategy relies on the assumption that treatment and control groups did not follow differential trends. Our dataset includes only one control cohort (2006) and therefore it's impossible to examine the existence of linear or non linear time trends. Nevertheless, as long as individual specific characteristics are time invariant, controlling for past performance would net out any factors that may be correlated with assignment in the treatment or the control group. Overall, we are of the view that the reform provides a valid instrument to identify returns to absences and peer effects.

One of the main worries it that eligibility is an endogenous variable. We do not include cohorts that graduated later than 2007 in the treatment group. These cohorts might exert more effort in the eleventh grade in order to exceed the performance threshold and become eligible. We use the graduating cohort of 2007 as the only treated group because the reform was announced after they had taken their eleventh grade final year exams.

The reform targeted only students that satisfied a specific criterion. However, this selection was based on a completely observable characteristic that determined eligibility. We control for the potential endogeneity of the eligibility by including the eleventh grade gpa in the main specification. The eligibility variable takes the value zero or one and as a non-linear function of the lagged GPA will not be perfectly collinear with lagged GPA.

Another worry would be that some schools misreport absences or the eleventh grade gpa. Teachers have to connect to a central electronic platform and upload the number of hours that a student is absent from class each month. If we believe that teachers report school attendance honestly in the first place, it will be difficult to change these numbers in a later month. Teachers have no incentives to misreport a student's school attendance or gpa. A concern would be that teachers in private schools might receive pressure from parents to misreport attendance or gpa. Private school teachers are not assigned to schools based on geographical criteria but the school administration decides the hiring practices. Although we dont believe that this could happen, one might be concerned that private school teachers might receive pressure to misreport students attendance or gpa. For these reasons, we excluded private schools from our analysis.

One may also be concerned that some principals/teachers were not satisfied by the reform's rationale and they refused to allow high achieving students to take the additional absences without a penalty. That would be a concern if there were many students exceeding the allowance. In fact, the reform allowed some students to skip school if they wanted but it did not force them to do so. In that sense the effect of the reform on the actual number of absences taken is the intention to treat effect (ITT). The average number of hours of total absences used in the twelfth grade is 76 (as we observe in Table 4.1). Even for students in the very right tail of the ability distribution the constraint is not binding ie. they take on average 80 hours of total absences (Figure 4.1). This alleviates another concern: teachers might react to the reform because high achieving students might be more absent from class so their classes might become nosier. This could happen because low achieving students tend to be more disruptive. That could affect a teacher's productivity in class as they might become less enthusiastic and motivated. As a consequence, we might observe a decrease in remaining students score that is not coming from their own school attendance. However, the first stage estimates that we find are not large enough to affect a teacher's motivation and behaviour in class.

We also use the reduced form regression that is essentially a difference-in-difference estimate of the reform:

$$Scores_{icstg} = \gamma_1 + \gamma_2 Eligibility_{icst} + \gamma_3 Reform_t + \gamma_4 Eligibility_{icst} * Reform_t$$

$$+studentFE + schoolFE + \epsilon_{icstg} (5)$$

We regress a student's score on the instrument proposed above. It compares changes in students' scores from the pre to the- post-reform period for eligible students to the change between the two years for non-eligible students. If the reform has a positive effect on the productivity of eligible students we would expect positive coefficients on the interaction term variable.

4.5.2 IV Results

The reform examined in this paper relaxed the attendance requirements of higher performing students and allowed them to skip more hours of class. This context offers itself to identification of the effect of class attendance on exam performance. The reform provides an exogenous source of variation. This is the eligibility status. The students considered in our study had no anticipation of the new absences law. Although the rationale behind the new law was rather to provide non pecuniary incentives to exert higher effort, in the short run it permits us to identify returns to absences.

Using both eligible and non-eligible students, we regress a student's standardised score in greek Language, mathematics and the twelfth grade gpa on the instrument proposed above. For students that are either non-eligible or in a year where the new law is not in place, the interaction term $Reform * Eligibility$ will take the value 0. Tables 4.6, 4.7 and 4.8 report the first stage, second stage and reduced form estimates.

In all regression we add students and school effects and we cluster the standard errors at the class level. Tables 4.6, 4.7 and 4.8 report the estimates for the following subjects respectively: Greek Language, mathematics and gpa. The first stage estimates imply that the reform increased total absences of eligible students by around 15 hours compared to non-eligible. This increase in the total absences in likely coming from students in the very right tail of the ability distribution. Notice that for students just to the right of the eligibility performance threshold we did not find a significant effect in the hours of absences taken. The reform is a strong and highly significant predictor of total absences. Our high F-statistics keep fears of weak instruments at bay.

Our reduced form estimates are also negative. The reform has a strong negative effect on eligible students' score that is always significant at 1%. This is the case when we use as outcome variables a student's performance in greek language, mathematics and the twelfth grade gpa. These results strongly support the view that the reform can be used as a valid source of exogenous variation in total absences.

Then, we investigate the effect of hours of total absences on a student's national exam performance using the regression model discussed above. Using the reform as an IV we overcome the problem that the total absences variable is endogenous. Tables 4.6, 4.7 and 4.8 report also the second stage estimates. We provide evidence of negative returns to absences in all subjects. In particular, an additional hour of absence implies a decrease by 0.012, 0.014 and 0.014 of a standard deviation in a student's national exam performance in Greek Language, Mathematics and gpa respectively. This is approximately similar to our benchmark OLS specification result (Table 4.5) for greek language, mathematics and gpa, but the IV results are almost six times larger than its corresponding benchmark (-0.002 compared to -0.012). This is of particular interest because these students may decide to be more absent from class in order to prepare for the senior-year exams. They might think that self-regulated learning helps them more than attending classes. We find evidence that skipping class even when a student chooses when to do so deteriorates a student's performance.

4.6 Heterogeneous Effects

We find that the returns to absences are negative in both subjects used (greek language and mathematics) and in the twelfth grade gpa. It might be helpful to investigate if the effects that we find differ based on some characteristics. Understanding any possible heterogeneity of the effects of an additional hour of absence might help the policy discussion by identifying groups of the population that are likely to disproportionately benefit from particular interventions.

In this section, we examine heterogeneity based on two dimensions: a student's gender and the size of the class. We run the same regressions as before separately for females and males. We use the reform as an instrumental variable and we report first stage and second stage estimates. In these specifications we control for student, school and grade fixed effects and we cluster standard errors at the class level. The IV estimates are reported in Table 4.9,4.10 and 4.11 for the different dependent variables: national exam performance in greek language, mathematics and twelfth grade gpa. The first stage estimates are higher for males compared to females. This means that boys are more absent from class by around 2.5 hours when they are allowed to do so. However, the effect that each additional hour has on their standardised score in greek language, mathematics and twelfth grade gpa is smaller compared to girls. In particular, an additional hour of absence from class is associated with a loss of 0.010 and 0.014 of a standard deviation in a student's national exam performance in greek language for boys and girls respectively. The difference between the two second stage estimates decline for mathematics. In particular, an additional hour of absence from class is associated with a loss of 0.013 and 0.014 of a standard deviation in a student's national exam performance in mathematics for boys and girls respectively. In terms of twelfth grade gpa, an additional hour of absence from class is associated with a loss of 0.011 and 0.017 of a standard deviation in a student's gpa for boys and girls respectively.

Then we explore the relationship between student attendance and performance for small and big classes separately. In Table 4.12, 4.13 and 4.14 we report the first and second stage estimates separately for students in small and big classes. We categorise a class as a small one if there are less than 15 students in this class. Otherwise it is considered as big. Our first stage estimates imply that students are on average 1 hour more absent from a big compared to a small class. Classes that consist of more than 15 students might be noisier and thus high-achieving students might choose to be more absent for 1 more hour when the reform is implemented. Although the difference between the first stage estimates are relatively small between small and big classes, this is not the case for the returns to absences. An additional hour of absence decreases a student's score in greek by 0.009 and 0.025 of a standard deviation in big and small classes respectively. Nevertheless, the size of the loss is larger in big classes in mathematics and gpa compared to small classes.

Future research is required to understand what is the mechanism that drives these heterogeneous effects. In particular, we need to understand why an additional day of absence from school hurts females more than it does males but also why the size of the loss in terms of math score may be larger in smaller classrooms, while the size of the loss in terms of greek language score may be larger in larger classrooms.

4.7 Conclusion

In this paper, we investigate returns to absences using a natural experiment. We exploit an unexpected reform that took place in Greece in 2007, that provided higher performing students with 50 more hours of excused absences from school. The eligibility status was determined based on a cut-off rule. We start off by using a Regression Discontinuity approach in order to measure the change in total absences and exam score due to the reform. Although, no strong effects were observed around the cut-off, important controls like class size and postcode income do not remain unchanged around the cut-off. This violates necessary assumption for identification in the regression discontinuity framework.

Next, we employ a combination of differences-in-differences and instrumental variables techniques in order to identify returns to absences. An interaction between eligibility status and year dummy is proposed as an instrument for the endogenous variable of total absences, to mitigate identification threats like unobserved heterogeneity.

Our study is the first one to identify returns to absences using a quasi-experimental approach. Our findings imply that students take on average 15 more hours of absence. Our estimates yield significant negative returns to absences in greek language, mathematics and twelfth grade gpa. Our results suggest that attendance is an important driver of school performance.

Figure 4.1: Hours of twelfth grade absences per term by prior performance

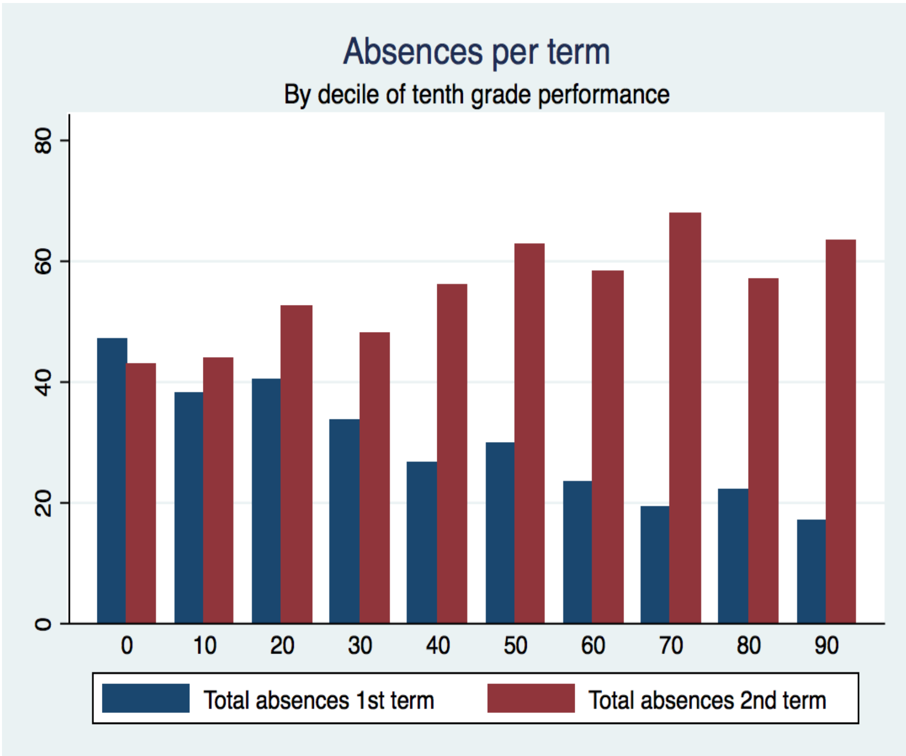


Figure 4.2: Regression Discontinuity Figures for Controls

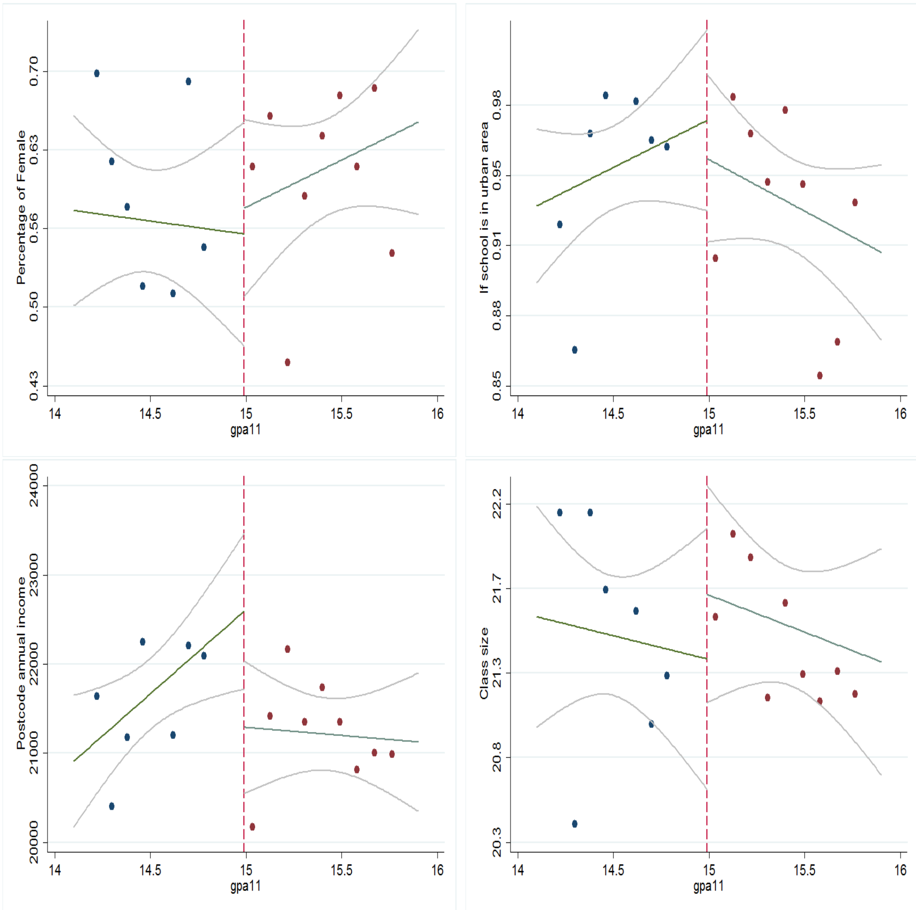


Figure 4.3: Regression Discontinuity Figures for First Stages with Different Bandwidths

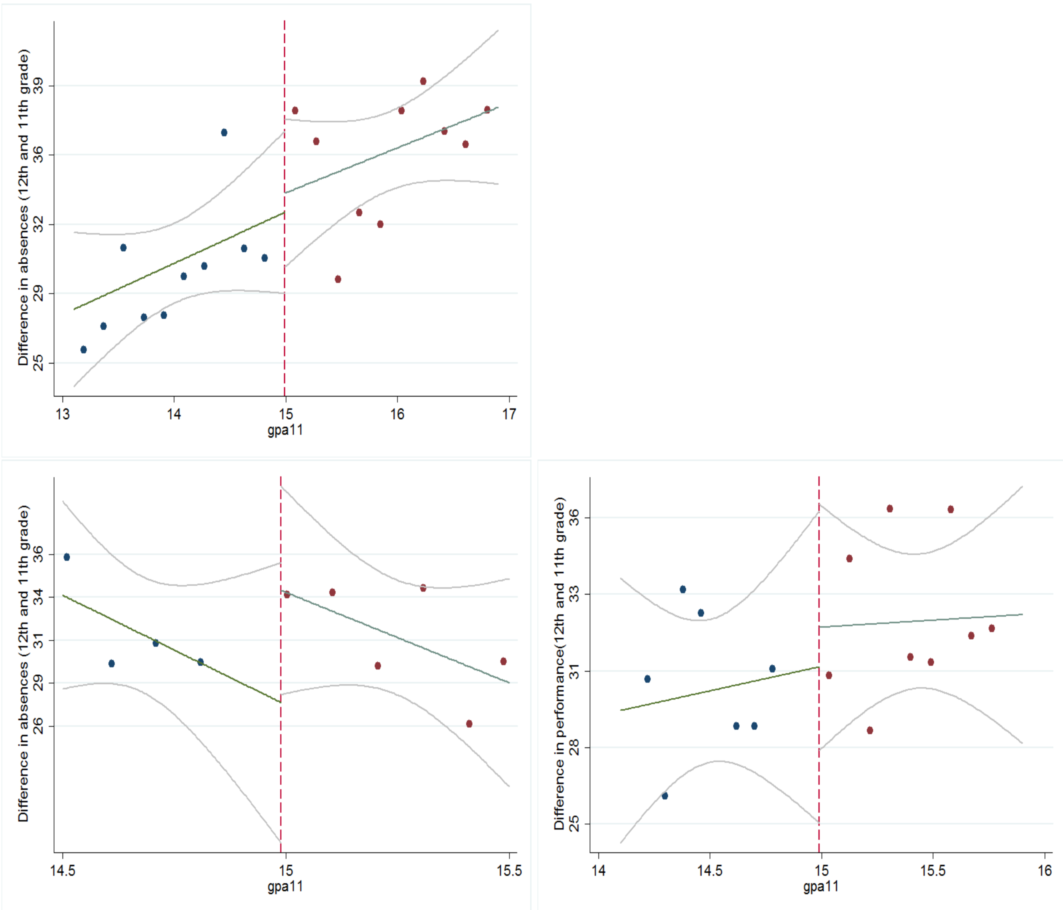


Table 4.1: Summary Statistics for full sample

Full Sample				
Variable	**Mean**	**Std. Dev.**	**Min**	**Max**
Born in 1st quarter	0.166	0.33	0	1
12th Grade Greek Language Score	12.50	3.06	0	20
11th Grade Greek Language Score	13.89	3.20	0	20
12th Grade Mathematics Score	10.85	6.39	0	20
11th Grade Mathematics Score	9.90	6.09	0	20
Female	0.57	0.50	0	1
Income (2009 Euro)	22,244	6,322	11,785	48,427
Experimental School	0.05	0.21	0	1
Public School	0.95	0.21	0	1
Urban	0.95	0.21	0	1
11th Grade GPA	14.21	2.85	8.8	20
12th Grade GPA	14.89	2.64	4.9	20
11th Grade Excused Absences	19.54	19.38	0	137
11th Grade Total Absences	49.53	27.38	0	164
12th Grade Excused Absences	42.01	23.91	0	160
12th Grade Total Absences	76.31	28.68	0	371

Note: sample: 11,239 obs.

Table 4.2: Summary Statistics

Variable	**Treated**	**Control**	**Diff**	**Std. Dev.**
Early Enrolment	0.10	0.20	0.11***	0.00
Female	0.56	0.57	0.01	0.01
Income (2009 Euro)	22,284	22,255	29.12	96.82
Experimental School	0.04	0.05	0.01	0.02
Public School	0.96	0.95	-0.01	0.01
Urban	0.95	0.95	0.00	0.00
11th Grade GPA	14.18	14.16	-0.07	0.04
12th Grade GPA	15.01	14.55	-0.46***	0.04
12th Grade Excused Absences	43.88	36.50	-7.38***	0.36
12th Grade Total Absences	78.22	70.68	-7.55***	0.44

Note: sample: 11,239 obs.

Table 4.3: Externalities in the classroom: Regression Discontinuity Estimates

Independent variable: GPA in 11th Grade	Non-Parametric approach					
	0.5/20	1/20	1.5/20	CCT(2014a)	IK (2012)	LM (2007)
Total Absences	6.648	2.236	2.226	1.503	1.360	1.740
	(6.254)	(4.064)	(3.214)	(3.065)	(2.031)	(1.773)
Mean in Control Group	69.026					
Observations	620	1,189	1,758	1,923	4,239	5,451
Δ(Total Absences)	0.051	1.714	2.335	-1.962	-1.879	-0.974
	(5.855)	(3.828)	(3.015)	(2.540)	(2.102)	(1.664)
Mean in Control Group	24.131					
Observations	593	1,111	1,689	2,288	3,281	5,527

Note: Δ(Total Absences) is the difference in total absences of a student between 11th and 12th grade. The estimation is conducted using a local linear regression constructed with a triangular kernel regression. Each column corresponds to a different bandwidth selection. Bias-corrected standard errors in parentheses. Column 1: 0.5 out of 20, Column 2: 1 out of 20, Column 3: 1.5 out of 20, Column 4: Calonico, Cattaneo and Titiunik (2014a), Column 5: Imbens and Kalyanaraman (2012), Column 6: Ludwig and Miller (2007). * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table 4.4: Externalities in the classroom: Regression Discontinuity Estimates

Independent variable: GPA in 11th Grade	Non-Parametric approach					
	0.5/20	1/20	1.5/20	CCT (2014a)	IK (2012)	LM (2007)
Greek Language	0.089	0.145*	0.139*	0.066	0.082*	0.079
	(0.135)	(0.075)	(0.071)	(0.075)	(0.042)	(0.050)
Observations	593	1,141	1,689	1,572	2,559	3,070
Δ(Greek Language)	0.193	0.197*	0.160*	0.077	0.071	0.084*
	(0.165)	(0.111)	(0.090)	(0.077)	(0.060)	(0.048)
Observations	593	1,141	1,689	2,175	3,702	5,270
Panel B						
Mathematics	-0.307*	-0.188*	-0.167*	-0.078	-0.070	-0.087*
	(0.169)	(0.103)	(0.092)	(0.074)	(0.067)	(0.050)
Observations	593	1,141	1,689	2,499	3,070	5,270
Δ(Mathematics)	-0.215	-0.343***	-0.286***	-0.175**	-0.136**	-0.066
	(0.161)	(0.102)	(0.082)	(0.089)	(0.063)	(0.049)
Observations	593	1,137	1,689	1,413	2,715	4,467
Panel C						
Grade Point Average	0.067	0.058	0.025	0.034	0.024	0.042
	(0.076)	(0.046)	(0.040)	(0.034)	(0.032)	(0.025)
Observations	620	1,189	1,758	2,436	2,542	4,075
Δ(Grade Point Average)	0.067	0.060	0.031	0.039	0.027	0.044*
	(0.075)	(0.049)	(0.041)	(0.034)	(0.033)	(0.024)
Observations	620	1,189	1,689	2,388	2,499	3,938

Note: Δ(Mathematics) is the difference in standardized score in Mathematics of a student between 11th and 12th grade. The estimation is conducted using a local linear regression constructed with a triangular kernel regression. Each column corresponds to a different bandwidth selection. Bias-corrected standard errors in parentheses. Column 1: 0.5 out of 20, Column 2: 1 out of 20, Column 3: 1.5 out of 20, Column 4: Calonico, Cattaneo and Titiunik (2014a), Column 5: Imbens and Kalyanaraman (2012), Column 6: Ludwig and Miller (2007). * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table 4.5: Naive Regression: OLS estimates

	OLS		
Variable	**Greek language**	**Mathematics**	**gpa**
Total Absences	-0.002	-0.001	-0.002
	(0.0003)***	(0.0002)***	(0.002)***
Senior year	-0.118	-0.242	-0.072***
	(0.027)***	(0.017)***	(0.012)***
School FE	✓	✓	✓
Student FE	✓	✓	✓

Sample: 11,238 obs. Standard Errors clustered at class level (377 clusters)

Table 4.6: IV Estimates: Greek Language

	First Stage	**Reduced Form**
Variable	**Total Absences**	**Greek**
*Reform*Eligibility*	14.883	-0.168
	(1.131)***	(0.029)***
Reform	-1.249	0.063
	(1.456)	(0.043)
Eligibility	0.434	-0.021
	(0.879)	(0.022)
Senior year	23.602	0.019
	(0.990)***	(0.032)
School FE	✓	✓
Student FE	✓	✓
Total Absences		-0.012
		(0.002)***
F-Statistic	173.08	

Sample: 11,238 obs. Standard Errors clustered at class level (377 clusters)

Table 4.7: IV Estimates: Mathematics

	First Stage	**Reduced Form**
Variable	**Total Absences**	**Mathematics**
*Reform*Eligibility*	14.883	-0.205
	(1.131)***	(0.022)***
Reform	-1.249	0.066
	(1.456)	(0.024)***
Eligibility	0.434	-0.049
	(0.879)	(0.015)***
Senior year	23.602	0.005
	(0.990)***	(0.016)
School FE	✓	✓
Student FE	✓	✓
		Second Stage
Total Absences		-0.014
		(0.002)***
F-Statistic	173.08	

Sample: 11,238 obs. Standard Errors clustered at class level (377 clusters)

Table 4.8: IV Estimates: Grade Point Average

	First Stage	**Reduced Form**
Variable	**Total Absences**	**GPA**
*Reform*Eligibility*	14.883	-0.205
	(1.131)***	(0.012)***
Reform	-1.249	0.075
	(1.456)	(0.017)***
Eligibility	0.434	-0.025
	(0.879)	(0.012)**
Senior year	23.602	0.058
	(0.990)***	(0.013)***
School FE	✓	✓
Student FE	✓	✓
		Second Stage
Total Absences		-0.014
		(0.001)***
F-Statistic	173.08	

Sample: 11,238 obs. Standard Errors clustered at class level (377 clusters)

Table 4.9: Males vs Females: Greek Language

	Females	**Males**
	First Stage	**First Stage**
Variable	**Total Absences**	**Total Absences**
*Reform*Eligibility*	13.704	16.423
	(1.352)***	(1.523)***
Reform	-1.788	-0.278
	(1.751)	(1.455)
Eligibility	1.155	-0.795
	(1.160)	(1.164)
Senior year	25.093	21.524
	(1.131)***	(0.984)***
F-Statistic	102.69	116.34
School FE	✓	✓
Student FE	✓	✓
	Second Stage	**Second Stage**
Total Absences	-0.014	-0.010
	(0.003)***	(0.003)***

Sample: 6,426 Females & 4,812 Males.

Table 4.10: Males vs Females: Mathematics

	Females	**Males**
	First Stage	**First Stage**
Variable	**Total Absences**	**Total Absences**
*Reform*Eligibility*	13.704	16.423
	(1.352)***	(1.523)***
Reform	-1.788	-0.278
	(1.751)	(1.455)
Eligibility	1.155	-0.795
	(1.160)	(1.164)
Senior year	25.093	21.524
	(1.131)***	(0.984)***
F-Statistic	102.69	116.34
School FE	✓	✓
Student FE	✓	✓
	Second Stage	**Second Stage**
Total Absences	-0.014	-0.013
	(0.002)***	(0.002)***

Sample: 6,426 Females & 4,812 Males.

Table 4.11: Males vs Females: Grade Point Average

	Females	**Males**
	First Stage	**First Stage**
Variable	**Total Absences**	**Total Absences**
*Reform*Eligibility*	13.704	16.423
	(1.352)***	(1.523)***
Reform	-1.788	-0.278
	(1.751)	(1.455)
Eligibility	1.155	-0.795
	(1.160)	(1.164)
Senior year	25.093	21.524
	(1.131)***	(0.984)***
F-Statistic	102.69	116.34
School FE	✓	✓
Student FE	✓	✓
	Second Stage	**Second Stage**
Total Absences	-0.017	-0.011
	(0.002)***	(0.001)***

Sample: 6,426 Females & 4,812 Males.

Table 4.12: Big vs Small Classrooms: Greek Language

	Small	**Big**
	First Stage	**First Stage**
Variable	**Total Absences**	**Total Absences**
*Reform*Eligibility*	17.887	18.691
	(4.929)***	(1.550)***
Reform	-2.192	1.069
	(4.415)	(2.708)
Eligibility	-8.315	-10.702
	(2.075)***	(0.757)***
Senior year	20.000	23.646
	(3.081)***	(2.074)***
F-Statistic	13.17	145.40
School FE	✓	✓
Student FE	✓	✓
	Second Stage	**Second Stage**
Total Absences	-0.025	-0.009
	(0.011)**	(0.003)***

Sample: Small if size<15, Big if size>15

Table 4.13: Big vs Small Classrooms: Mathematics

	Small	Big
	First Stage	**First Stage**
Variable	**Total Absences**	**Total Absences**
*Reform*Eligibility*	17.887	18.690
	(4.929)***	(1.550)***
Reform	-2.192	1.069
	(4.415)	(2.708)
Eligibility	-8.315	-10.702
	(2.075)***	(0.757)***
Senior year	20.000	23.646
	(3.081)***	(2.074)***
F-Statistic	13.17	145.40
School FE	✓	✓
Student FE	✓	✓
	Second Stage	**Second Stage**
Total Absences	-0.008	-0.014
	(0.006)	(0.003)***

Sample: Small if size<15, Big if size>15

Table 4.14: Big vs Small Classrooms: Grade Point Average

	Small	Big
	First Stage	**First Stage**
Variable	**Total Absences**	**Total Absences**
*Reform*Eligibility*	15.687	16.449
	(1.352)***	(1.472)***
Reform	-0.861	1.910
	(4.460)	(2.709)
Eligibility	-0.424	-2.431
	(3.271)	(2.057)
Senior year	20.000	24.230
	(3.055)***	(2.062)***
F-Statistic	10.45	124.91
School FE	✓	✓
Student FE	✓	✓
	Second Stage	**Second Stage**
Total Absences	-0.011	-0.016
	(0.006)**	(0.002)***

Sample:Small if size<15, Big if size>15

Bibliography

Aaronson, D., Barrow, L., and Sander, W. (2007). Teachers and Student Achievement in the Chicago Public High Schools. *Quarterly Journal of Economics*, 25(1):95–135.

Angrist, D. J. and Lavy, V. (1999). Using Maimonides'Rule to Estimate the Effect of Clas Size on Scholastic Achievement. *Quarterly Journal of Economics*, 114(114):533–575.

Angrist, J. D. and Krueger, A. B. (1992). The Effect of Age at School Entry on Educational Attainment: An Application of Instrumental Variables with Moments from Two Samples. *Journal of the American Statistical Association*, 87(418):328–336.

Arcidiacono, P. and Nicholson, S. (2005). Peer effects in medical school. *Journal of Public Economics*, 89(23):327–350.

Armin, F. and Andrea, I. (2006). Clean evidence on peer effects. *Journal of Labor Economics*, 24(1).

Arulampalam, W., Naylor, R., and Smith, J. (2012). Am i missing something? *Economics of Education Review*, 31(4):pp. 363–375.

Azmat, G. and Iriberri, N. (2010). The importance of relative performance feedback information: Evidence from a natural experiment using high school students. *Journal of Public Economics*, 94(5):797–811.

Bandiera, O., Larcinese, V., and Rasul, I. (2008). Blissful Ignorance? Evidence from a Natural Experiment on The Effect of Individual Feedback on Performance. *Working Paper*.

Barankay, I. (2012). Evidence from a Randomized Workplace Experiment. *Management Science, Working Paper*.

Barry J., Z. and Manuel, M. P. (1990). Student differences in self regulated learning: Relating grade, sex, and giftedness to self efficacy and strategy use. *Journal of Educational Psychology*, 82(1):pp. 51–59.

Bell, A. P. (1970). Role modelship and interaction in adolescence and young adulthood. *Developmental Psychology*, 2.

Bernheim, B. D. (1994). A Theory of Conformity. *Journal of Political Economy*, 102(5):841–877.

Brock, W. A. and Durlauf, S. N. (2001). Discrete Choice with Social Interactions. *The Review of Economic Studies*, 68(2):235–260.

Bryk, A. and Raudenbush, S. (2001). Hierarchical linear models, second edition. Technical report.

Calonico, S., Cattaneo, M., and Titiunik, R. (2014). Robust data driven inference in the regression discontinuity design. *The Stata Journal*, 14(4):pp. 909–946.

Card, D. and Krueger, B. A. (1992). Does School Quality Matter? Returns to Education and the Characteristics of Public Schools in the United States. *The Journal of Political Economy*, 100(1):1–40.

Caviglia Harris, J. L. (2006). Attendance and achievement in economics: Investigating the impact of attendance policies and absentee rates on student performance. *Journal of Economics and Finance education*, 4(2):pp. 470–476.

Chen, J. and Lin, T.-F. (2008). Class Attendance and Exam Performance: A Randomized Experiment. *The Journal of Economic Education*, 39(3):213–227.

Chetty, R., Friedman, J., and Rockoff, J. (2014). Measuring the Impacts of Teachers I: Evaluating Bias in Teacher Value-Added Estimates. *American Economics Review*, 104(9):2593 2632.

Chong, T., Cheung, K.-S., and Hui, P.-H. (2009). Skipping economics classes: a case study from hong kong. *Journal of Higher Education Policy and Management*, 31(1):pp. 37–42.

Cipollone, P. and Alfonso, R. (2007). Social interactions in high school: Lessons from an earthquake. *American Economic Review*, 97(3):948–965.

Cohn, E. and Johnson, E. (2006). Class Attendance and Performance in Principles of Economics. *Education Economics*, 14(2):211–233.

Ertac, S. (2005). Social Comparisons and Optimal Information Revelation:Theory and Experiments,. *mimeo, University of California.*

Fletcher, J. (2006). Social interactions in college choice: The interplay of information, preferences, social norms and individual characteristics in predicting college choice. *Working Paper.*

Franz, T., Frick, E., and Hanslits, K. (2009). Gender Differences in Response to Failure Feedback. *Working paper.*

Garner, C. L. and Raudenbush, S. W. (1991). Neighborhood Effects on Educational Attainment: A Multilevel Analysis. *Sociology of Education*, 64(4):251–262.

Giorgi, G. D., Pellizzari, M., and Redaelli, S. (2007). Be as careful of the books you read as of the company you keep: Evidence on peer effects in educational choices. *NBER Working Papers 14948*, pages 241–275.

Hahn, J., Todd, P., and der Klaauw, W. V. (2001). Identification and estimation of treatment effects with a regression-discontinuity design. *Econometrica*, 69(1):pp. 201–209.

Hannan, R. L., Krishnan, R., and Newman, A. H. (2008). The Effects of Disseminating Relative Performance Feedback in Tournament and Individual Performance Compensation Plans. *The Accounting Review*, 83(4):893–913.

Hanushek, A. E., Kain, F. J., and Rivkin, G. S. (2005). Teachers, Schools and Academic Achievement. *Econometrica*, 73(2):417–458.

Hanushek, E., Kain, J., Markman, J., and Rivkin, S. (2003). Does peer ability affect student achievement? *Journal of Applied Economics*, 18(5):527–544.

Heckman, J. (2001). Micro Data, Heterogeneity, and the Evaluation of Public Policy: Nobel Lecture. *Journal of Political Economy*, 109(4):673–748.

Hoxby, C. (2000a). Peer Effects in the Classroom: Learning from Gender and Race Variation. Working Paper 7867, National Bureau of Economic Research.

Hoxby, C. M. (2000b). Peer Effects in the Classroom: Learning from Gender and Race Variation. *NBER Working Paper*, 7867.

Imbens, G. and Kalyanaraman, K. (2012). Optimal bandwidth choice for the regression discontinuity estimator. *Review of Economic Studies*, 79(3):pp. 933–959.

Krueger, B. (1999). Experimental Estimates of Education Production Functions. *Quarterly Journal of Economics*, 114(114):497–532.

Latif, E. and Miles, S. (2013). Class attendance and academic performance: A panel data analysis. *Economics Papers*, 32(4):pp. 470–476.

Lavy, V. and Schlosser, A. (2011). Mechanisms and Impacts of Gender Peer Effects at School. *American Economic Journal: Applied Economics*, 3(2):1–33.

Lavy, V., Silva, O., and Weindhardt, F. (2012). The Good, The Bad and the Average: Evidence on Ability Peer Effects in Schools. *Journal of Labor Economics*, 20(12):367–414.

Levine, J. (April 3-5, 1992). The effect of different attendance policies on student attendance and achievement, paper presented at the annual meeting of the eastern psychological association boston, ma.

Lin, T.-F. and Chen, J. (2006). Cumulative class attendance and exam performance. *Applied Economics Letters*, 13(14):937–942.

Ludwig, J. and Miller, D. (2007). Does head start improve childrens life chances? evidence from a regression discontinuity design. *Quarterly Journal of Economics*, 122(1):pp. 159–208.

Manski, C. F. (1993). Identification of Endogenous Social Effects: The Reflection Problem. *The Review of Economic Studies*, 60(3):531–542.

Manski, C. F. (2000). Economic Analysis of Social Interactions. Working Paper 7580, National Bureau of Economic Research.

McCarty, P. A. (1986). Effects of Feedback on the Self-Confidence of Men and Women. *Journal of Academy of Management*, 29(4):840–847.

Moore, R. (2006). Class attendance: How students' attitudes about attendance relate to their academic performance in introductory science classes. *Research and Teaching in Developmental Education*, 23(1):pp. 19–33.

Murphy, R. and Weinhardt, F. (2014). Top of the class: The Importance of Rank Position. *London School of Economics and Political Science, Job Market Paper.*

Nickell, S. and Van Reenen, J. (2001). Technological innovation and performance in the united kingdom. *Working Paper CEP*, (0488).

Nicola, D. and Debra, M. D. (2006). Formative assessment and self regulated learning a model and seven principles of good feedback practice. *Studies in Higher Education*, 31(2):pp. 199–218.

Raudenbush, S. and Sampson, R. (1999). Ecometrics: Towards a science of assessing ecological settings, with application to the systematic social observation of neighborhoods. Technical report.

Romer, D. (1993). Do students go to class?should they? *The Journal of Economic Perspectives*, 7(3):pp. 167–174.

Rothstein, J. (2010). Teacher Quality in Educational Production: Tracking, Decay, and Student Achievement. *The Quarterly Journal of Economics*, 125(1):175–214.

Sacerdote, B. (2011). Peer Effects in Education: How Might They Work, How Big Are They and How Much Do We Know Thus Far? Handbook of the Economics of Education, Elsevier.

Schelling, T. C. (1973). Hockey Helmets, Concealed Weapons, and Daylight Saving: A Study of Binary Choices with Externalities. *The Journal of Conflict Resolution*, 17(3):381–428.

Thistlethwaite, D. L. and Campbell, D. T. (1960). Regression-discontinuity analysis: An alternative to the ex post facto experiment. *Journal of Educational Psychology*, 51(6):pp. 309–317.

Tremblay, K. (2005). Academic Mobility and Immigration. *Journal of Studies in International Education*, 9(3):196–228.

Zimmerman, D. J. (2003a). Peer Effects in Academic Outcomes: Evidence from a Natural Experiment. *Review of Economics and Statistics*, 85(1):9–23.

Zimmerman, D. J. (2003b). Peer Effects in Academic Outcomes: Evidence from a Natural Experiment. *Review of Economics and Statistics*, 85(1):9–23.